

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 4, 2002, 00:26:31 ; Search time 6206.58 Seconds  
(without alignments)  
8107.924 Million cell updates/sec

Title: US-08-956-991-1\_COPY\_453\_5135

Perfect score: 4683  
Sequence: 1 alytgagactgctctctc.....gctgcgcgagagcagc 4683

Scoring table: IDENTITY-NUC  
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*  
1: em\_estfun:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estcom:\*  
5: em\_estcpl:\*  
6: em\_estba:\*  
7: em\_estro:\*  
8: em\_estov:\*  
9: em\_hic:\*  
10: qd\_est1:\*  
11: qd\_est2:\*  
12: qd\_hic:\*  
13: qd\_gss:\*  
14: em\_gss\_fun:\*  
15: em\_gss\_hum:\*  
16: em\_gss\_inv:\*  
17: em\_gss\_pln:\*  
18: em\_gss\_pro:\*  
19: em\_gss\_rod:\*  
20: em\_gss\_vct:\*  
21: em\_gss\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIPS

Result No.	Score	Query Match	Length	DB ID	Description
1	382.4	12.4	653	11	BG819678
2	405.8	8.7	499	11	B1185857
3	348.8	7.4	455	11	BF544715
4	343.8	7.3	355	10	AJ003472
5	341.2	7.3	458	10	A1454704
6	282	6.0	487	10	BG187880
7	274.2	5.9	410	10	AM529324
8	271.2	5.8	427	10	AM159035
9	254.8	5.4	483	11	BG307959
10	254	5.2	744	11	BF348620
11	244	5.0	344	13	A2576100
12	234.8	5.0	1232	12	AK015320

13	234.8	5.0	1376	12	AK014832
14	228.4	4.9	468	11	BF565509
15	219.8	4.7	279	10	BB336040
16	217.6	4.6	291	10	BB333216
17	216.2	4.6	355	10	BB461534
18	202.4	4.3	252	10	BB381728
19	202.4	4.3	287	11	BF928859
20	202.2	4.3	270	10	BB336019
21	201	4.1	248	10	BB335137
22	191.2	4.1	266	10	BB329941
23	188.4	4.0	247	10	BB330789
24	188	4.0	280	10	BB331669
25	182.8	3.9	257	10	BB330426
26	179.6	3.8	610	13	AZ56759
27	179.4	3.8	253	10	BB331552
28	178.6	3.8	304	10	BR280255
29	175.6	3.7	619	13	FR0002941
30	172.6	3.7	286	10	BB333910
31	169.8	3.6	272	10	BB330770
32	169	3.6	252	10	BB334670
33	167.8	3.6	239	10	BB332000
34	163.2	3.5	402	10	AM048129
35	154	3.3	849	13	CNS031Y7
36	150.6	3.2	3133	12	CNS031Y6
37	149.2	3.2	216	10	BB328887
38	148.6	3.2	409	10	AM014717
39	147.4	3.1	219	10	BB342624
40	144.6	3.1	147	13	HSWC18B12
41	143	3.0	912	13	CNS021Y6
42	140.8	2.9	491	13	AZ657601
43	138	2.9	212	10	BB330864
44	136.2	2.9	967	13	CNS025E3
45	135.4	2.9			

#### ALIGNMENTS

RESULT 1  
LOCUS BG819678 653 bp mRNA EST 22-MAY-2001  
DEFINITION 602784172P1 NCI\_CGAP\_Brn67 Homo sapiens cDNA clone IMAGE:4934248  
5' mRNA sequence.  
ACCESSION BG819678  
KEYWORDS  
SOURCE  
ORGANISM Human.  
REFERENCE NIH-MGC http://mgi.nci.nih.gov/.  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE Unpublished (1998)  
JOURNAL Contact: Robert Strausberg, Ph.D.  
COMMENT Email: cgapbs-rt@mail.nih.gov  
Tissue Procurement: David N. Louis, M.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: L1AM10862 row: 0 column: 17  
High quality sequence stop: 648.  
Location/Qualifiers  
1. 653  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4934248"  
/clone\_lib="NCI\_CGAP\_Brn67"  
/tissue\_type="anaplastic oligodendroglioma with 1p/19q loss"

#### FEATURES

Source



[illegible]

Db	68	TGCTCTCTTTCACACACCCGCGAGTGTGCGCTTTTGAGACATATGGAGCTCGACACTGTACG	127
Qy	1688	ctgtgtcaaaaggaagatggagaaaggaggaagacagctgtgcaagctgtgtgtgttcaaccacaac	1747
Db	128	ATGTGCAGAAAGAAAGCTTGACATAGGAGGAGATACACATGTATGTGGTGCGAGCCAGC	187
Qy	1748	ctctccatcagctcagagcgtctccagctgaccgtgtgaagcttcgcgccttcatacaacccttg	1807
Db	188	TCTCCACCAAGCCAGATGCTTCACGTGACAGTCAAGTCTCTCTCTTATTCACACCTTTTG	247
Qy	1808	aattctccaaattctccatctgagcgaacgaattctatcccttgctgtgtgagctcaagg	1867
Db	248	AGTTCCGAGAGATCTCTATCTCCCTCAGCGGCTCTCATCCATCTGTGGTGCGTGCGGGG	307
Qy	1868	acttactccatcagaatcactctgagcagaagaaatggcgcgcgaatcccttggaagccttgg	1927
Db	308	ACTTACCTATCCATCCATCACTCCGAGAGATGGCCGGCGGATTCACGCAAGCCTTGCGTG	367
Qy	1928	tgacattctgacatatttgaactcagcagcctccttgagagattccaaatcctctgcgtcatgc	1987
Db	368	TAAACATTAATAAATGTCGACTTCAATCACTCTCTGTGAGAGATCTCCAACTCTACATGATGC	427
Qy	1988	acaatctgaaattatcaactgcaatgaagctgg	2015
Db	428	ACACGCGCAATTATACCTGTATGCGCGG	455

[illegible]







Fax: 516 367 8874  
Email: mcomble@cs.hi.org  
Plate: z450 row: e column: 05  
Seq primer: M13 universal forward primer  
High quality sequence stop: 427.  
Location/Qualifiers

## FEATURES

source

1. 427  
/organism="Xenopus laevis"  
/db\_xref="taxon:8355"  
/clone\_lib="z450e05"  
/clone\_lib="Xenopus EST library"  
/tissue\_type="total brain tissue"  
/cell\_line="W22-TGA"  
/dev\_stage="tadpole"  
/note="Vector: Lambda zap I; site\_1: XbaI; This library was supplied by Holly Cline (Cold Spring Harbor Labs). cDNA synthesis with oligo dt Xba I (Xba I cloning site). RNA: stage 50-56 tadpoles, total brain tissue, gtc extraction method."

3 COUNT 114 a 122 c 91 g 100 t

Query Match 5.8%; Score 271.2; DB 10; Length 427;

Best Local Similarity 79.8%; Pred. No. 2.3e-65; Indels 7; Gaps 1;

Matches 335; Conservative 0; Mismatches 78; Indels 7; Gaps 1;

1226 ttttttcgttcttgatgaagaagtgatgagtcacagagagccggttcccttattgca 1285

15 tcatctctgcatTTAGGAGAGAGTGTGAGCCCGAGAACCTGTCTCTATGTGCA 74

1286 acgtgaagaagaaccttggccagatcagtgagccctggagtgatccagatctca 1345

75 acgtgaagaagaaccttggccagatcagtgagccctggagtgatccagatctca 134

1346 aggggtgagcagtcacagcagtcacagtcacagtcacagtcacagtcacagtcac 1405

135 AGAAGCGCCACACAGATAAGTACGTTAACTCTGAGGCGAGTTGTGAGTTACG 194

1406 tgaacatctcagctccagcagtcagcagtcagcagtcagcagtcagcagtcagcag 1465

195 TGACATCACCACACATCAAGTCCGAGATGCGGGGTGTACCCGATGCGGCTPACACT 254

1466 cagtgagagatctgctcgtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 1525

255 CTGCGCGGCTCTCTTGTACCGAGCTCGAATTAACCT-----CGTCGAAGTATCCGTC 307

1526 caatgaanaacatcacagcaatagcagagcagcagcagcagcagcagcagcagcagc 1585

308 CAATGAANAACATCACAGCGCCATACGCTGCGCGACATCATCCAGTCGCGCTCATTTG 367

1586 gctatcgtatlaactcattaaatggtacagaaactctaacctgcttcttllaacacac 1645

368 GATATCCCTACTACTCCATCAAGTGTGACAAACCTTAACCTCTCCCTTATATACAC 427

## RESULTS

9

LOCUS BG307959 483 bp mRNA EST 22-FEB-2001

DEFINITION fms5b04.y1 zebrafish adult retina cDNA Danio rerio cDNA c. ne

MOLECULE: ., mRNA sequence.

ACCESSION BG307959 GI:13105486

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

Danio rerio  
zebrafish.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;  
Cypriniformes; Cyprinidae; Rasbora; Danio.

## REFERENCE

1 (bases 1 to 483)  
Clark, M., Johnson, S.L., Lehrbach, H., Lee, R., Li, F., Morris, M., Ruddy,  
S., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wyllie, T., Underwood

TITLE  
JOURNAL  
COMMENT

Washington University School of Medicine  
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: zbratfish@wustl.edu  
Library constructed by: Susan E. Brockerhoff DNA Sequencing by:  
Washington University Genome Sequencing Center Clone distribution:  
Resourcenzentrum Immundatenbank, Berlin, Germany (web address:  
www.rzpd.de)  
Seq primer: T3 RT from Amersham  
High quality sequence stop: 406.  
Location/Qualifiers

## FEATURES

source

1. 483  
/organism="Danio rerio"  
/strain="wild-type"  
/db\_xref="taxon:7955"  
/clone="419358"  
/clone\_lib="zebrafish adult retina cDNA"  
/sex="mixed"  
/dev\_stage="1-2 years"  
/lab\_host="E.Coli XL1-Blue MRF" (XL1-Blue MRF)  
/note="Vector: Lambda ZAP II (pBluescript SK-); Site\_1:  
EcoRI; Site\_2: SalI; This zebrafish library was  
constructed by Dr. Susan E. Brockerhoff (email:  
sbrocker@wustl.edu) RZPD library number: 760"

BASE COUNT 112 a 130 c 114 g 107 t

Query Match 5.4%; Score 254.8; DB 11; Length 483;

Best Local Similarity 73.2%; Pred. No. 1.1e-60; Indels 2; Gaps 2;

Matches 353; Conservative 0; Mismatches 127; Indels 2; Gaps 2;

3891 tccatgataaagaacatgcttctcgttgaagcgttgggaacctctcctgagc 3950

1 tccctgcatgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgac 60

3951 caaatggaatggaatggaatggaatggaatggaatggaatggaatggaatggaatgga 4010

61 CAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGA 120

4011 gacatcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 4070

121 CGCTCACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAA 180

4071 gacatcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 4130

181 CACATGCTGTCAC 240

4131 agtctcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 4188

241 AGTAC 300

4189 cttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 4248

301 CTCTTCTGATCCCGGACGACAAATGGCGGAAGCTCCATCGACGCGTACATTTTACAGTAC 360

4249 tccgagacaaatagtgagcagtgagcagtgagcagtgagcagtgagcagtgagcagtgag 4308

361 TCAGAGCATTAACACTGAAGATGGGAGTACTCCATGACGCCCAAGCGCTCTTATAC 420

4309 cgccttgaaatctcaaatgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 4368

421 CGTCGAGAAATTTGAGAGTCCGACTTGTGACAGATTACTCTGACTGACGAGAAATGCA 480

4369 gtc 4370





Db	770	gtagctgattccctcttccactgcagagaatracggttaccgcgtctcttgcgaataagacac	829
Db	492	gattcc 497	
Db	830	gggtctc 835	
RESULT	13		
LOCUS	AK014832		
DEFINITION	AK014832	1376 bp mRNA	05-JUL-2001
ACCESSION	AK014832	Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:4921507G06, full insert sequence.	
VERSION	AK014832.1	GI:12852903	
KEYWORDS	CAP trapper.	Mus musculus (strain:C57BL/6J) adult male testis cDNA to mRNA, clone:4921507G06.	
SOURCE			
ORGANISM		Mus musculus	
REFERENCE		Eukaryotes: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Murinae: Mus	
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			

[illegible]



prepared and sequenced in Mouse Genome Encyclopedia  
Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in  
RIKEN, Division of Experimental Animal Research in Riken  
contributed to prepare mouse tissues. 1st strand cDNA was  
primed with a primer [5',  
GAGAGAGAGAGATCCAGAGAGCTTTTCTTTTCTTTTCTTTT 3'], cDNA was  
prepared by using trehalose L: rmo-activated reverse  
transcriptase and subsequently enriched for full-length by  
cap-trapper. cDNA went through one round of normalization  
to RoL = 20.0 and subtraction to RoL = 459.0. Second  
strand cDNA was prepared with the primer adapter of  
sequence [5' GAGAGAGAGATTCGAGTAAATTAATCCCCCCCC  
3']. cDNA was cleaved with XhoI and BamHI. Vector: a  
modified phluescript KS(+) after bulk excision from lambda  
FLC I."

BASE COUNT 81 a 68 c 68 g 62 t  
ORIGIN

Query Match 4.7%; Score 219.8; DB 10; Length 279;  
Best Local Similarity 86.7%; Pred. No. 6.6e-51;  
Matches 242; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

OY 2488 cctgagatggccgttatctgtgtccaccaagaagtgagagaagatgattctact 2547  
DB 1 CTTGAATGGCCATCTACTGTTATCCACCATGGAGAGAGTGAATATACAG 60  
OY 2548 ctgagatttgcgaacttgagagaagatctggttcttctcctgacctatataat 2607  
DB 61 CTGCAATTTCCCAACACAGTGAAGAGATTCGGTTT ATTTCCTGCTCATGCTATCAAT 120  
OY 2608 tcttctggggagagccatggaataatcagctcagtgcaagagccccaagacctccc 2667  
DB 121 TCATACGGGGAGACCGTGGAAATTAATTCACAGTGAAGAACCCACAGATCTCC 180  
OY 2668 gaaattgagatcaaatgltcaagaagcagcaaatlaagctcaagtggaccatggggtt 2727  
DB 181 GAGATTGAGATCAAGAGTGTCAAGCTCGACCATCAAGTTCAGGTGAGCCATGGGCTT 240  
OY 2728 gatgaaacagtcctcatcacagagctagatattgaatgc 2766  
DB 241 GATGCCAACAGCCCATCAGGCTTATGACATGATGC 279

Search completed: March 4, 2002, 02:14:32  
Job time: 6481 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 4, 2002, 00:29:31 ; Search time 578.64 Seconds  
(without alignments)  
6938.432 Million cell updates/sec

Title: US-08-956-991-1\_COPY\_453\_5135  
Perfect score: 4683  
Sequence: 1 atgtgatactgtgtctctc.....gtctgcgagagacgagcc 4683

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: N\_Geneseq\_1101:.\*  
2: /SID52/gcgdata/geneseq/geneseqn/NA1980.DAT:.\*  
3: /SID52/gcgdata/geneseq/geneseqn/NA1981.DAT:.\*  
4: /SID52/gcgdata/geneseq/geneseqn/NA1982.DAT:.\*  
5: /SID52/gcgdata/geneseq/geneseqn/NA1983.DAT:.\*  
6: /SID52/gcgdata/geneseq/geneseqn/NA1984.DAT:.\*  
7: /SID52/gcgdata/geneseq/geneseqn/NA1985.DAT:.\*  
8: /SID52/gcgdata/geneseq/geneseqn/NA1986.DAT:.\*  
9: /SID52/gcgdata/geneseq/geneseqn/NA1987.DAT:.\*  
10: /SID52/gcgdata/geneseq/geneseqn/NA1988.DAT:.\*  
11: /SID52/gcgdata/geneseq/geneseqn/NA1989.DAT:.\*  
12: /SID52/gcgdata/geneseq/geneseqn/NA1990.DAT:.\*  
13: /SID52/gcgdata/geneseq/geneseqn/NA1991.DAT:.\*  
14: /SID52/gcgdata/geneseq/geneseqn/NA1992.DAT:.\*  
15: /SID52/gcgdata/geneseq/geneseqn/NA1993.DAT:.\*  
16: /SID52/gcgdata/geneseq/geneseqn/NA1994.DAT:.\*  
17: /SID52/gcgdata/geneseq/geneseqn/NA1995.DAT:.\*  
18: /SID52/gcgdata/geneseq/geneseqn/NA1996.DAT:.\*  
19: /SID52/gcgdata/geneseq/geneseqn/NA1997.DAT:.\*  
20: /SID52/gcgdata/geneseq/geneseqn/NA1998.DAT:.\*  
21: /SID52/gcgdata/geneseq/geneseqn/NA1999.DAT:.\*  
22: /SID52/gcgdata/geneseq/geneseqn/NA2000.DAT:.\*  
23: /SID52/gcgdata/geneseq/geneseqn/NA2001.DAT:.\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4683	100.0	6604	19	AAV31981 Human Down syndrom
2	4678.8	99.9	6413	22	AAV31988 Human Down syndrom
3	1806.8	38.6	5824	19	AAI58196 Human polynucleoti
4	636	13.6	842	19	AAV31985 Mouse Down syndrom
5	529.2	11.3	1493	19	AAV27207 cDNA clone etbb001
6	502	10.7	898	19	AAV31986 Mouse Down syndrom
7	299.4	6.4	2173	19	AAV31987 Human polynucleoti
8	227.6	4.9	471	22	AAI59982 Human polynucleoti
9	139	3.0	388	19	AAV31982 Human Down syndrom
10	91.4	2.0	3617	22	AAH24799 Nucleotide sequenc
11	91.4	2.0	4336	22	AAH05447 Human secreted pro

12	78	1.7	808	22	AAH06899 Human cDNA clone (
13	78	1.7	2537	22	AAH14602 Human PTP-OR. Hom
14	66.4	1.4	6000	16	AAO86478 Human protein tyro
15	66.4	1.4	6000	18	AAH85389 Human protein tyro
16	66.4	1.4	6000	20	AAH06095 Human protein tyro
17	66.4	1.4	6000	22	AAH02144 Rat hypothetical 1
18	66.2	1.4	2401	22	AAH01017 Human protein tyro
19	57	1.2	7702	21	AAH8739 Human cDNA encodin
20	57	1.2	7702	21	AAZ59133 Human cDNA encodin
21	57	1.2	7705	22	AAH98405 Human EST-derived
22	57	1.2	7705	22	AAH52684 Human cDNA encodin
23	57	1.2	7741	22	AAH22448 Drosophila sp. ROB
24	57	1.2	7741	20	AAH57250 Rat oploid recepto
25	53.2	1.1	4188	20	AAH57250 Rat oploid recepto
26	53.2	1.1	2179	14	AAH051017 Rat oploid recepto
27	49.6	1.1	2337	14	AAH051016 Mechanical stress
28	49.6	1.1	306	14	AAH051015 Mechanical stress
29	49.6	1.1	837	21	AAH26301 Mechanical stress
30	48	1.0	7679	21	AAH26322 Mechanical stress
31	48	1.0	8882	21	AAH26321 Human LI cell adhe
32	48	1.0	3774	22	AAH01598 Rat receptor type-
33	47.8	1.0	3774	22	AAH01598 Human UNC-40 nucle
34	46.4	1.0	7710	22	AAH44662 Deleted in colorec
35	45.8	1.0	5690	16	AAH086902 Human DOC DNA. Ho
36	45.2	1.0	4344	22	AAH090961 Human folistatin-
37	45.2	1.0	4608	16	AAH080196 Human polynucleoti
38	45.2	1.0	4608	20	AAH23431 Human polynucleoti
39	45.2	1.0	5186	12	AAH12752 Human polynucleoti
40	45.2	1.0	2526	22	AAH30917 Human polynucleoti
41	44.8	1.0	3730	22	AAH159396 Human polynucleoti
42	44.6	1.0	3783	22	AAH161182 Human polynucleoti
43	44.6	1.0	4843	22	AAH37553 Human polynucleoti
44	44.6	1.0	5202	22	AAH99637 Human protein enco
45	44.2	0.9	5994	15	AAH05674 Sequence encoding

## ALIGNMENTS

RESULT 1  
AAV31981 standard: cDNA: 6604 bp.  
ID AAV31981:  
XX AAV31981:  
AC 28-SEP-1998 (first entry)  
XX  
XX Human Down syndrome: cell adhesion molecule DS-CAM) cDNA.  
DE  
XX US-CAM1: Down syndrome: cell adhesion molecule: neural cell:  
XX signal transduction: fibrocyte 21: mental retardation:  
XX holoprosencephaly: corpus callosum agenesis:  
XX schizencephaly: diagnosis: assay: human: ds: ss.  
XX  
XX Homo sapiens.  
OS  
XX  
XX Key Location/Qualifiers  
FH 453..6185  
FT /\*tag- a  
FT sig-peptide 452..521  
FT /\*tag- b  
FT mat-peptide 522..6162  
FT /\*tag- c  
XX  
XX W09817795-A1.  
XX  
XX 30-APR-1998.  
XX  
XX 23-OCT-1997: 97WO-US19547.  
XX  
XX 25-OCT-1996: 96US-0029322.  
XX  
XX (CEDA-) CINDARS SINAI MEDICAL CENT.



xx Korenberg JR;  
xx  
xx WPI: 1998-271791/24.  
xx  
xx P-PSDB: AAM42086.  
xx  
xx  
xx New isolated Down's Syndrome-cell adhesion molecule - used to  
xx develop products for detection, diagnosis and therapy of  
xx developmental and neurological abnormalities  
xx  
xx Claim 3; Page 65-72; 109pp; English.

This cDNA clone codes for Down syndrome-cell adhesion molecule DS-CAM1 (see AAM42086), a cell surface glycoprotein belonging to a novel subclass of the Ig superfamily with highest homology to neural cell adhesion molecules. A modified direct cDNA selection technique was applied to bacterial and P1 artificial chromosome libraries between ET52 and MX1 by using cDNA from trisomy 21 human foetal brain. A unique cDNA fragment, designated E51 (see AAV31982), was obtained and used to screen a trisomy 21 human foetal brain (14-W gestation) cDNA library. Further clones were obtained by exon trapping. The complete DS-CAM1 cDNA sequence was deduced from overlapping clones. A splice variant cDNA (see AAV31988), encoding non-membrane bound DS-CAM2 (see AAM42087), was also identified. The DS-CAM1 gene spans 900-1200 kb of genomic DNA and maps at chromosome 21q22.2-22.3. The invention also provides murine DS-CAM partial sequences (see AAV31985-87), expression vectors and host cells, transgenic animals, antisense oligonucleotides, and primers useful for amplification of DS-CAM nucleic acid. DS-CAM polypeptides are associated with developmental and neurological processes. They can be used in e.g. neural prosthetic devices used in encephalation methods of repairing (regenerating) damaged or severed peripheral nerves. The products can also be used in detection, diagnosis and therapy of developmental and neurological abnormalities such as Down syndrome, mental retardation, holoprosencephaly, agenesis of the corpus callosum, or schizencephaly. Antisense oligonucleotides are used for inhibiting translation of mRNA.

Sequence 6604 BP; 1673 A; 1827 C; 1768 G; 1336 T; 0 other;

Query Match	100.0%	Score 4683	DR 19	Length 6604
Best local Similarity	100.0%	Pred. No. 0		
Matches 4683: Conservative	0	Mismatches	0	Indels 0
				Gaps 0

[illegible]

OY	421	gcgcgtctccaagtgcattalcccctctccgtgtagagcgtagacatcaactgtctctcaatg	480
Db	873	gcgcgtctccaagtgcattalcccctctccgtgtagagcgtagacatcaactgtctctcaatg	932
OY	481	gaagaagaacacttttcaattttctttccaggatctaaattttccatcaatccacggggacc	540
Db	933	gaagaagaacacttttcaattttctttccaggatctaaattttccatcaatccacggggacc	992
OY	541	ttgcatcttaaaagatgacagaaatgaaagatgtaattgtaataactccgtgcataccggtg	600
Db	993	ttgcatcttaaaagatgtaagaaatgaaagatgtaattgtaataactccgtgcataccggtg	1052
OY	601	catctgatattcccttgagagatgaaatgcaagaaacaaacggccgaactttttgatacggacca	660
Db	1053	catctgatattcccttgagagatgaaatgcaagaaacaaacggccgaactttttgatacggacca	1112
OY	661	ggcgaactctcagcccatcccatctcagatctgagatcttgagccatctgcaagaaccatgctggagcag	720
Db	1113	ggcgaactctcagcccatcccatctcagatctgagatcttgagccatctgcaagaaccatgctggagcag	1172
OY	721	cgtgtctgagagctgtcccttgagaaatgctctgggaccccttgagccagatataccgtctgctaaag	780
Db	1173	cgtgtctgagagctgtcccttgagaaatgctctgggaccccttgagccagatataccgtctgctaaag	1232
OY	781	gacacatacgcctccctggagaaacttcaggagagatctccgaagaacccgtatcaagggtctgtcatc	840
Db	1223	gacacatacgcctccctggagaaacttcaggagagatctccgaagaacccgtatcaagggtctgtcatc	1292
OY	841	gagaaacatctgcgcctctcgaactctcaggccagcatatgtttgttgaagtgatccacaagaatacggga	900
Db	1293	gagaaacatctgcgcctctcgaactctcaggccagcatatgtttgttgaagtgatccacaagaatacggga	1352
OY	901	actgtttaaagtgtaaaagctccctctgtagctgtagaaacagccactgtaaaagccacatcatgctcc	960
Db	1353	actgtttaaagtgtaaaagctccctctgtagctgtagaaacagccactgtaaaagccacatcatgctcc	1412
OY	961	aggaaaggtctlaaaagcagatgtagatctgagccaaagtcttctgtctcagcgctgaacgggaact	1020
Db	1413	aggaaaggtctlaaaagcagatgtagatctgagccaaagtcttctgtctcagcgctgaacgggaact	1472
OY	1021	gagagaccagagaaactctctcttgtagccgaatggtggaatactctcaaacctctgtaaaaaatgctg	1080
Db	1473	gagagaccagagaaactctctcttgtagccgaatggtggaatactctcaaacctctgtaaaaaatgctg	1532
OY	1081	aaagatctaaatgagatctaaatctaaatgaaactcttattaaatgaaatctgtaaaagtgaacggg	1140
Db	1533	aaagatctaaatgagatctaaatctaaatgaaactcttattaaatgaaatctgtaaaagtgaacggg	1592
OY	1141	gagcgtatctccctttgcttattgttttaagatctaaatggttctccgcctcaagatctatgctcaaggtg	1200
Db	1593	gagcgtatctccctttgcttattgttttaagatctaaatggttctccgcctcaagatctatgctcaaggtg	1652
OY	1201	gtccctctgaagaatggaactctcctaaadaatlatctctgcgctcttaagtgaagaagtgtagtgcctca	1260
Db	1653	gtccctctgaagaatggaactctcctaaadaatlatctctgcgctcttaagtgaagaagtgtagtgcctca	1712
OY	1261	gtcagaagcttgcttctccctctatgtagcaagctgaagggaaacacactcttgccacgatacagctggg	1320
Db	1713	gtcagaagcttgcttctccctctatgtagcaagctgaagggaaacacactcttgccacgatacagctggg	1772
OY	1321	accctctgaagaatgtagcccggaattcttcaaaagtgtagaagctcacgcgatacagccagatgtagtcag	1380
Db	1773	accctctgaagaatgtagcccggaattcttcaaaagtgtagaagctcacgcgatacagccagatgtagtcag	1832
OY	1381	tcggagagggaaacgttgtagcagctgaacatctccagctctcccaaggtctcggggaacgggggga	1440
Db	1833	tcggagagggaaacgttgtagcagctgaacatctccagctctcccaaggtctcggggaacgggggga	1892
OY	1441	gtctcaacgctctcaactgccaacaacatctggggggagctgcgtctcttaccagagctctgaaataaac	1500
Db	1893	gtctcaacgctctgactgccaacaacatctggggggagctgcgtctcttaccagagctctgaaataaac	1952
OY	1501	gttaaaagagctctgcaaaagatcttcgaaccatgtaaaaaatcacacgaataagcagagcggagc	1560

[illegible]

Db	3033	ggtttctcttctccctgcacatgctatataattctctatctggtggagagacgttggaataattcagctc	3092
Oy	2641	acagtgcaagagccccccagacccctcccgaaatctgagatccaagaatgttcaaaagcagcgaca	2700
Db	3093	acagtgcaagagccccccagacccctcccgaaatctgagatccaagaatgttcaaaagcagcgaca	3152
Oy	2701	attacgcttcaggtgagacacatggtgttgatggaaacagttcccatccacggttaagatatt	2760
Db	3153	attacgcttcaggtgagacacatggtgttgatggaaacagttcccatccacggttaagatatt	3212
Oy	2761	gaattgcaaaataaatacgaactccctggagattctgtctcagagaaaccaagaatgtttccctt	2820
Db	3213	gaattgcaaaataaatacgaactccctggagattctgtctcagagaaaccaagaatgtttccctt	3272
Oy	2821	cagcttgatattcgtgcacacattcatgtatattccacccctccctccacacatcagatccgcattg	2880
Db	3273	cagcttgatattcgtgcacacattcatgtatattccacccctccctccacacatcagatccgcattg	3332
Oy	2881	tacgccaagaacccggaatttgccaagagcgagcccgacaacagagctaccatccacgcgagac	2940
Db	3333	tacgccaagaacccggaatttgccaagagcgagcccgacaacagagctaccatccacgcgagac	3392
Oy	2941	gaagcagctcctctgattggttccacacctcaagaaagttccaccttgagacctatcatctcaagac	3000
Db	3393	gaagcagctcctctgattggttccacacctcaagaaagttccaccttgagacctatcatctcaagac	3452
Oy	3001	atcacggttccacttggaaagcttcccaagaacacttggcaaatggagattatctcgtgtgctac	3060
Db	3453	atcacggttccacttggaaagcttcccaagaacacttggcaaatggagattatctcgtgtgctac	3512
Oy	3061	caaatatgatttaccggaagatatacgcattggggatlaacttccaattccaatatcatctcagtttc	3120
Db	3513	caaatatgatttaccggaagatatacgcattggggatlaacttccaattccaatatcatctcagtttc	3572
Oy	3121	gaacacacagtggtgacagctgaagtttatacccttcggacaacctgaatgaattgactacgaac	3180
Db	3573	gaacacacagtggtgacagctgaagtttatacccttcggacaacctgaatgaattgactacgaac	3632
Oy	3181	ggcctctggtgattgctcagcgtctgtaaacccggcgccgacgagggcctctctctcagaagaatcattc	3240
Db	3633	ggcctctggtgattgctcagcgtctgtaaacccggcgccgacgagggcctctctctcagaagaatcattc	3692
Oy	3241	accacacacctctcgagagatctctgcccagttacccccccggaaaattctcagaacgcatagaacaa	3300
Db	3693	accacacacctctcgagagatctctgcccagttacccccccggaaaattctcagaacgcatagaacaa	3752
Oy	3301	taccacaaagaacatattcgaattctctgattctcgaattcttccaaagaacgttcaaatggaaatt	3360
Db	3753	taccacaaagaacatattcgaattctctgattctcgaattcttccaaagaacgttcaaatggaaatt	3812
Oy	3361	ctctgaagatttccatgatttcatctggtggtcaattctcatlygacgtagagctcgggttgagatt	3420
Db	3813	ctctgaagatttccatgatttcatctggtggtcaattctcatlygacgtagagctcgggttgagatt	3872
Oy	3421	aaaaaaccttaccac	3480
Db	3873	aaaaaaccttaccac	3932
Oy	3481	tacacgcatctccagtgctctggtcttccaccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc	3540
Db	3933	tacacgcatctccagtgctctggtcttccaccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc	3992
Oy	3541	ttcaccccgatccaaagagagatgttccaggtctctcccgcggtgtgtaagagcagcgcgcgcc	3600
Db	3993	ttcaccccgatccaaagagagatgttccaggtctctcccgcggtgtgtaagagcagcgcgcgcc	4052
Oy	3601	tcagccttccatctgctcttctggtctccggtcttccacctctcaaaccttgaaacgcgcgcgcgcgcgc	3660
Db	4053	tcagccttccatctgctcttctggtctccggtcttccacctctcaaaccttgaaacgcgcgcgcgcgcgc	4112
Oy	3661	aagtaacatctatattctctctccacccctatacccaagatgatacagcgagatttgagagctctct	3720
Db	4113	aagtaacatctatattctctctccacccctatacccaagatgatacagcgagatttgagagctctct	4172

```
QY 3721 ccgcactcgtttccctacagaattcccaacctgaatgagaaatcgtcagctacagcgtctg 3780
    |||
DB 4173 ccgcactcgtttccctacagaattcccaacctgaatgagaaatcgtcagcgtctg 4232
QY 3781 gtagtgagcgttacttcagccggaagaagcacaacagcagatcatcatcacaatgcagca 3840
    |||
DB 4233 gtagtgagcgttacttcagccggaagaagcacaacagcagatcatcatcagctcagccca 4292
QY 3841 ctacgcaaaagctctcagcagcaatcctgacctccagtgaggacagtgactctccatggatg 3900
    |||
DB 4293 ctacgcaaaagctctcagcagcaatcctgacctccagtgaggacagtgactctccatggatg 4352
QY 3901 aaagacatcgtctcgtctgtatgagcgtcgttctggaacctctccctccatcgaatcgaatg 3960
    |||
DB 4353 aaagacatcgtctcgtctgtatgagcgtcgttctggaacctctccctccatcgaatcgaatg 4412
QY 3961 aaagacatcgtctcgtctgtatgagcgtcgttctggaacctctccctccatcgaatcgaatg 4020
    |||
DB 4413 aaagacatcgtctcgtctgtatgagcgtcgttctggaacctctccctccatcgaatcgaatg 4472
QY 4021 aacggaagcttcatctcagcagcgtggaagcagaagcagcagctatcagcagctcagc 4080
    |||
DB 4473 aacggaagcttcatctcagcagcgtggaagcagaagcagcagctatcagcagctcagc 4532
QY 4473 aacggaagcttcatctcagcagcgtggaagcagaagcagcagctatcagcagctcagc 4532
    |||
DB 4081 gccataacaactcgggagcagctcgtatgaaatcattttaaactatacagtaacagttccacca 4140
    |||
DB 4533 gccataacaactcgggagcagctcgtatgaaatcattttaaactatacagtaacagttccacca 4592
QY 4141 gatcagcctcgggttactagatcctcccaagacacagctctccctccatccctctcttggtc 4200
    |||
DB 4593 gatcagcctcgggttactagatcctcccaagacacagctctccctccatccctctcttggtc 4652
QY 4201 ccttgagacacaacggggcagcagctctacacagagatatactcagtaactcccgaggaat 4260
    |||
DB 4653 ccttgagacacaacggggcagcagctctacacagagatatactcagtaactcccgaggaat 4712
QY 4261 agtgagcagtgaggagcttctccaaatcagcccgagcgaagcttccctctcttggtcgaat 4320
    |||
DB 4713 agtgagcagtgaggagcttctccaaatcagcccgagcgaagcttccctctcttggtcgaat 4772
QY 4321 ctcaaatctgagcagcttggtatgaatctcaacatcagcagcccaaaatgagagtgagccag 4380
    |||
DB 4773 ctcaaatctgagcagcttggtatgaatctcaacatcagcagcccaaaatgagagtgagccag 4832
QY 4381 cgcataaatctgaatcatalcatalgaagcaagacctctgaaaaagcgcctccatctcuaaag 4440
    |||
DB 4833 cgcataaatctgaatcatalcatalgaagcaagacctctgaaaaagcgcctccatctcuaaag 4892
QY 4441 caggagcctgcttctccagcagcctcaacacacagcgtctgagacgtctcatttgctggaat 4500
    |||
DB 4893 caggagcctgcttctccagcagcctcaacacacagcgtctgagacgtctcatttgctggaat 4952
QY 4501 gatgagcagctcgcctccatcactcctctcagcagtaagtaacagcctcttgagacacagct 4560
    |||
DB 4953 gatgagcagctcgcctccatcactcctctcagcagtaagtaacagcctcttgagacacagct 5012
QY 4561 tggaccacacagctcagagagacctctctctccaaagtcactcactatgataatctgagagaa 4620
    |||
DB 5013 tggaccacacagctcagagagacctctctctccaaagtcactcactatgataatctgagagaa 5072
QY 4621 gccaccctgataagagctcagagtgaggctgagcaacagcagcagcagcagcagcagcagcag 4680
    |||
DB 5073 gccaccctgataagagctcagagtgaggctgagcaacagcagcagcagcagcagcagcagcag 5132
QY 4681 gcc 4683
    |||
DB 5133 gcc 5135
```

RESULT 2  
AAV31988  
ID AAV31988 standard: cDNA: 6413 BP.

```
XX AAV31988:
AC 28-SEP-1998 (first entry)
XX
DE Human Down syndrome-cell adhesion molecule DS-CAM2 cDNA.
XX
KW DS-CAM2; Down syndrome-cell adhesion molecule; neural cell;
KW signal transduction; trisomy 21; mental retardation;
KW holoprosencephaly; corpus callosum agenesis;
KW schizencephaly; diagnosis; assay; human; ds; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Journal
FT CDS 453..5168
FT /*tag= a
XX
PN M09817795-A1.
XX
PD 30-APR-1998.
XX
XX 23-OCT-1997; 97WD-US19547.
XX
XX 25-OCT-1996; 96US-0029322.
XX
PA (CEDA-) CEDARS SINAI MEDICAL CENT.
XX
PI Korenberg JR.
XX
DR WPI: 1998-271791/24.
DR P-PSDB: AAM42087.
XX
PT New isolated Down's Syndrome-cell adhesion molecule - used to
PT develop products for detection, diagnosis and therapy of
PT developmental and neurological abnormalities
XX
PS Claim 3: Page 83-90; 1 pp: English.
XX
XX This cDNA clone codes for Down syndrome-cell adhesion molecule
CC DS-CAM2 (see AAM42087), an extracellular, soluble protein belonging
CC to a novel subclass of the Ig superfamily with highest homology to
CC neural cell adhesion molecules. A modified direct cDNA selection
CC technique was applied to bacterial and P1 artificial chromosomes
CC between EFS2 and MX1 by using cDNA from trisomy 21 human foetal
CC brain. A unique cDNA fragment, designated E51 (see AAV31982), was
CC obtained and used to screen a trisomy 21 human foetal brain (14-wk
CC gestation) cDNA library. Positive clones were obtained by exon
CC trapping, and the complete DS-CAM2 cDNA sequence was identified.
CC DS-CAM2 is a splice variant of membrane-bound DS-CAM1 (see AAM42086)
CC and is deleted for the entire transmembrane domain. The DS-CAM
CC gene spans 900-1200 kb of genomic DNA and maps at chromosome
CC 21q22.2-22.3. The invention also provides murine DS-CAM partial
CC sequences (see AAV31985-87), expression vectors and host cells,
CC transgenic animals, antisense oligonucleotides, and primers useful
CC for amplification of DS-CAM nucleic acid. DS-CAM polypeptides are
CC associated with developmental and neurological processes. They can
CC be used in e.g. neural prosthetic devices used in entubulation
CC methods of repairing (regenerating) damaged or severed peripheral
CC nerves. The products can also be used in detection, diagnosis and
CC therapy of developmental and neurological abnormalities such as
CC Down syndrome, mental retardation, holoprosencephaly, agenesis of
CC the corpus callosum, or schizencephaly. Antisense oligonucleotides
CC are used for inhibiting translation of mRNA.
XX
SQ Sequence 6413 BP: 1633 A; 1781 C; 1707 G; 1292 T; 0 other;
```

Query Match 99.98; Score 4678.8; DB 19; Length 6413;  
Best local similarity 100.0%; Pred. No. 0;  
Matches 4680; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 atgttgatcgtctctctctgtctcagagcgttcgagatgtttcagtgagacctta 60







[illegible][illegible]



[illegible]

Db	4196	gcatcgctctctgagtcacgcgcccacaaggagacctgagcctgg	---	caggcctccggcc	4252
QY	4582	tctctctcacaagtcctacatcctctatgacgcacaagcaacactggtatgagctcag			4661
Db	4253	aacagctcccgaggagtgcttctctgacaggaactcgcgagagccacgttgtagagactgcgc			4312
QY	4642	atgcgcgtctgacaacagctgcgcggctgcg			4669
Db	4313	atgaagccttccacaacatgcgcggctgcg			4340

## RESULT 4

ID AAV31985 standard; cDNA; 842 BP.

AAV31985;

28-SEP-1998 (first entry)

XX  
DE Mouse Down syndrome-cell adhesion molecule DS-CAM 5' cDNA.

XX  
XX  
NS-CAM: Down syndrome; cell adhesion molecule; neural cell;

signal transduction; LI150my 21; mental retardation; KIM

KW schizencephaly; diagnosis; assay; mouse; ds; ss-

OS Mus sp.

PN W09817795-A).

30-APR-1998.

23-OCT-1997; 97WU-US19547.

25-OCT-1996: 96US-0029322

XX  
DA  
(CEDA-) CEDARS STNAT  
-JUDICAL, CENT

XX  
XX  
Responsible IIIXX  
1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  

XX  
XX

...the ... molecule - used to

PT develop products for detection, diagnosis and therapy of

XX  
XX

23 21 20 19 18 17 16 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1  
 XX

for murine Brown syngeneic cell adhesion molecule (DS-CAM),

member of a novel subclass of the  $\alpha$ -glycosaminoglycan family. The middle region (see AA

CC and 3' region (see AAV.3.1987) of the clone are also present in the 3' region of the clone isolated from a C57 B/Black/6 mouse.

CC cDNA library using human DS-CAM cDNA (see AAV31901 and AAV31902).

CC (AAW42086-87), as well as expression vectors and host cells

for simplification of DS-CAM nucleic acids. DS-CAM polypeptides

be used in e.g. neural prosthetic devices used in entubula

The products can also be used in detection, diagnosis

Down syndrome, mental retardation, holoprosencephaly, agen-

are used for inhibiting translation of mRNA.

XX 643 DM: 193 A: 244 C: 238 G: 178 T: 0 other:

SQ Sequence 842 BP; 182 A; 244 C; 238 G; 178 T; 0 Occurrences

DB 19: Tennant 842:

Query Match

13.68; Score 636; DB 19; Length 842;





[illegible]

ID	AAV31986 standard; cDNA; 898 BP.	
XX		
AC	AAV31986:	
XX		
DT	28-SEP-1998 (first entry)	
XX		
DE	Mouse Down syndrome-cell adhesion molecule DS-CAM mid cDNA.	
XX		
KM	DS-CAM; Down syndrome-cell adhesion molecule; neural cell;	
XX	signal transduction; trisomy 21; mental retardation;	
KM	holoprosencephaly; corpus callosum agenesis;	
XX	schizencephaly; diagnosis; assay; mouse; ds; ss.	
OS		
XX	Mus sp.	
XX		
PN	W09817795-A1.	
PD	30-Apr-1998.	
XX		
PF	23-OCT-1997; 97MO-US19547.	
XX		
PR	25-OCT-1996; 96US-0029322.	
XX		
XX	(CEDA-) CEDARS SINAI MEDICAL CENT.	
XX		
PI	Korenberg JK;	
DR	WPI; 1998-271791/24.	
XX		
PT	New isolated Down's Syndrome-cell adhesion molecule - used to	
PT	develop products for detection, diagnosis and therapy of	
PT	developmental and neurological abnormalities	
XX		
XX	Claim 2; Page 81; 109pp; English.	
CC	This cDNA sequence comprises the middle region of a cDNA clone	
CC	for murine Down syndrome-cell adhesion molecule (DS-CAM), a	
CC	member of a novel subclass of the 19 superfamily with homology to	
CC	neural cell adhesion molecules. The 5' region (see AAV31985)	
CC	and 3' region (see AAV31987) of the clone are also provided. The	
CC	murine DS-CAM clone was isolated from a C57 Black/6 mouse brain	
CC	cDNA library using human DS-CAM cDNA (see AAV31981 and AAV31988) as	
CC	probe. The invention also provides human DS-CAM proteins (see	
CC	AAW2086-87), as well as expression vectors and host cells.	
CC	transgenic animals, antisense oligonucleotides, and primers useful	
CC	for amplification of DS-CAM nucleic acids. DS-CAM polypeptides are	
CC	associated with developmental and neurological processes. They can	
CC	be used in e.g. neural prosthetic devices used in entubulation	
CC	methods of repeating (rejuvenating) damaged or severed peripheral	
CC	nerves. The products can also be used in detection, diagnosis and	
CC	therapy of developmental and neurological abnormalities such as	
CC	Down syndrome, mental retardation, holoprosencephaly, agenesis of	
CC	the corpus callosum, or schizencephaly. Antisense oligonucleotides	
CC	are used for inhibiting translation of mRNA.	
XX		
SO	Sequence 898 BP; 174 A; 222 C; 245 G; 257 T; 0 other;	
	Query Match 10.7%; Score 502; DB 19; Length 898;	
	Best Local Similarity 82.3%; Pred. No. 6,5e-136;	
	Matches 749; Conservative 0; Mismatches 135; Indels 26; Gaps 14;	
QY	2308 aacgatgtggccaga--cgtagcaagtcctgtaccccaagttaaattcctgcga 2365	
DB	892 AACGATGTGGCGGAGAACGTTTCAGCAGTTCATGCACTCACAGGGAA-----TTCCGG 838	
QY	2366 tgaataatcctaatcaataactacccttgcgcacagagggcagaaaagagatgagct 2425	
DB	837 CCAATGATATTACTATATACCCACACACACCTGCACATCGGCTCAAGAGAGA-ATGAGCT 779	
QY	2426 gcaacgcacatggtgaaagaccatataatcgtgagagagagagagaccgaatcata 2485	
DB	778 GCACAGGCCCAAG-GGAGAGCCCATTCATTGTTCGCT-GGAGAGAGAGACAGCATCATTA 722	



RESULT 8  
ID AAI59982 standard; cDNA: 471 BP.  
XX  
AC AAI59982:  
XX  
DT 22-OCT-2001 (first entry)  
DE Human polynucleotide SEQ ID NO 3971.  
XX  
KW Human; neotropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia; ss.  
XX  
KW Homo sapiens.  
OS  
MO200153312-AL.  
PD 26-JUL-2001.  
PF 26-DEC-2000; 2000MO-US34263.  
XX  
PR 21-JAN-2000; 2000US-0488725.  
PR 25-APR-2000; 2000US-0552317.  
PR 09-JUL-2000; 2000US-0598042.  
PR 19-JUL-2000; 2000US-0620312.  
PR 03-AUG-2000; 2000US-0653450.  
PR 14-SEP-2000; 2000US-0662191.  
PR 19-OCT-2000; 2000US-0693036.  
PR 29-NOV-2000; 2000US-0727344.  
XX  
XX (HSE-) HSEQ INC.  
PA Tang YF, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ken F, Wang D;  
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
PI Zhao QA, Zhou P, Goddard R, Drmanac RT;  
XX  
DR WPI: 2001-442253/47.  
DR P-PSDB: AAM40826.  
XX  
XX Novel nucleic acids and polypeptides, useful for treating disorders  
XX such as central nervous system injuries -  
XX  
XX Claim 1: SEQ ID NO 3971: 10078bp: English.  
XX  
XX The invention relates to human nucleic acids (AA157798-AA161369) and  
XX the encoded polypeptides (AAM38642-AAM42213) with neotropic.  
XX CC Immunosuppressant and cytostatic activity. The polynucleotides are useful  
XX in gene therapy. A composition containing a polypeptide or polynucleotide  
XX of the invention may be used to treat diseases of the peripheral nervous  
XX system, such as peripheral nervous injuries, peripheral neuropathy and  
XX localised neuropathies and central nervous system diseases, such as  
XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
XX utilisation of the activities such as: immune system suppression,  
XX activating/inhibiting activity, chemotactic/chemokinetic activity, haemostatic  
XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
XX assays for receptor activity, arthritis and inflammation, leukaemias and  
XX C.N.S disorders.  
XX Note: the sequence data for this patent did not form part of the printed  
XX specification.  
XX  
XX Sequence 471 BP: 96 A: 154 C: 130 G: 91 T: 0 other:  
XX  
XX Query Match 4.9%; Score 227.6; DB 22; Length 471;  
XX Best Local Similarity 69.0%; Pred. No. 5,6e-56;  
XX Matches 327; Conservative 0; Mismatches 144; Indels 3; Gaps 1;

QY 1610 ggtacaagaacttcaaccctgtctccttcaaccaccgcaagtgagcattttagaacaatg 1669  
DB 1 ggtacaagaactgtccctgtcgtccagacaaccaccgcaagtgagtggtttagaaca---tg 57  
QY 1670 gaactcttaaaccttcaagatgttgcaaaaggaagtgagcagagggagtagacagtgcaag 1729  
DB 58 ggaacctcaagctgaagctgaagctgaagaaagggcattgagatgagggagtagactgtgca 117  
QY 1730 tgttggttcaaccacacatcttccaccagcagcagcgttcaagctgaacgttgaagtccgc 1769  
DB 118 tccctcatcagccaccagcttccatcagcagcagcgttcaagctgaacgttgaagtccgc 177  
QY 1790 ctcttcaacaaccccttcttgatcttcaagatcttcttggcagcggtcttcaaccct 1849  
DB 178 ctcttcaacaaccccttcttgatcttcaagatcttcttggcagcggtcttcaaccct 237  
QY 1850 gttgt 1909  
DB 238 gttgt 297  
QY 1910 tcccttggaggt 1969  
DB 298 tcccttggaggt 357  
QY 1970 ccaacttctgt 2029  
DB 358 ctgagctgt 417  
QY 2030 ctgt 2083  
DB 418 ctgt 471  
RESULT 9  
AAV31982/c  
ID AAV31982 standard; cDNA: 388 BP.  
XX  
XX AAV31982:  
AC 28-SEP-1998 (first entry)  
XX  
XX Human Down syndrome-cell adhesion molecule DS-CAM probe E51.  
XX  
XX DS-CAM1: Down syndrome-cell adhesion molecule; neural cell;  
XX signal transduction; trisomy 21; mental retardation;  
XX holoprosencephaly; corpus callosum agenesis;  
XX schizencephaly; diagnosis; assay; human; ds; ss.  
XX  
XX Homo sapiens.  
OS  
XX  
XX W09817795-A1.  
XX  
XX 30-APR-1998.  
XX  
XX 23-OCT-1997; 97MO-US19547.  
XX  
XX 25-OCT-1996; 96US-0029322.  
XX  
XX (CEDA-) CEPRAKS SINAI MEDICAL CENT.  
XX  
XX Korenberg JR;  
XX  
XX WPI: 1998-271791/24.  
XX  
XX New isolated Down's Syndrome-cell adhesion molecule - used to  
XX develop products for detection, diagnosis and therapy of  
XX developmental and neurological abnormalities  
XX  
XX Example 2: Page 79: 109pp; English.  
XX  
XX This polynucleotide comprises cDNA fragment E51 that was isolated  
XX from a human trisomy 21 foetal brain (14-wk gestation) cDNA  
XX library following a modified direct cDNA selection technique

CC applied to bacterial and P1 artificial chromosomes between EMS2  
CC and MX1. ESI was used as a probe to screen the trisomy 21 foetal  
CC brain library. 62 Clones were isolated from the 2 million clones  
CC of the original library. Overlapping clones were obtained that  
CC were used to deduce a full-length sequence (see AAV31981) coding  
CC for novel Down syndrome-cell adhesion molecule DS-CAM2 (see  
CC AAM42086). A splice variant, DS-CAM2 (see AAV31988), was also  
CC identified. DS-CAM polypeptides are associated with developmental  
CC and neurological processes. The polypeptides and nucleic acids  
CC are used to develop products for the detection, diagnosis and  
CC therapy of developmental and neurological abnormalities.

XX Sequence 388 BP: 83 A: 119 C: 95 G: 91 T: 0 other:

Query Match 3.0%: Score 139; DB 19; Length 388;  
Best Local Similarity 100.0%: Pred. No. 3.6e-30;  
Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

4545 ctctggagccacagcttggaccacagctcagaagacctctctccaaagctccatccct 4604  
|||||  
388 CTTTGCGACACAGCTTGACACAGCTCAGAGCAGCTCTCTCCAAAGCTTCAATCTCT 329  
OY 4605 gtatgacrtgaggaagcaccctgtgtatgactgacatgacgtgtg:gcacaaatgcggg 4664  
|||||  
nb 328 GTATGACCTGACGAGAGCAGCTGTATGACCTGACATGCGGTGTCCAAAGCTGCGCG 269  
OY 4665 ctgcgcggagagacagcagcc 4683  
|||||  
nb 268 CTGCGCGAGAGAGAGCGCC 250  
RESULT 10  
AAH24799  
ID AAH24799 standard; DNA: 3617 BP.  
XX  
AC AAH24799;  
XX  
DT 06-AUG-2001 (first entry)  
XX  
DE Nucleotide sequence of a human cytokine receptor-like protein.  
XX  
KW Igg Fc region; cytokine receptor-like protein; immune system disorder;  
KW inflammatory disorder; immune modulation; immune cell stimulation;  
KW allergy; cancer; cardiovascular disorder; ankyrogenic disorder;  
KW wound healing; infectious disease; neural disorder; SS.  
XX  
Homo sapiens.  
Key Location/Qualifiers  
FT CDS 2..2968  
FT /tag= a  
FT /product= "cytokine receptor-like protein"  
XX  
PN MO200140456-A1.  
XX  
PD 07-JUN-2001.  
XX  
PF 30-NOV-2000; 2000MO-US32525.  
XX  
PR 03-DEC-1999; 99US-0168621.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Ruben SM, Ni J, Young PE, Shi Y;  
XX  
DR WPI: 2001-367808/38.  
XX  
DR P-PSDB: AAB84252.  
XX  
XX Nucleic acids encoding 2 human cytokine receptor-like polypeptides,  
XX useful for preventing, diagnosing and/or treating e.g., cancers,  
XX cardiovascular and immune disorders -

PS Claim 1; Page 237-238; 247pp; English).

XX The present sequence encodes a human cytokine receptor-like protein.  
XX The cytokine receptor-like polypeptides and polynucleotides may be used  
XX in the prevention, diagnosis and treatment of diseases associated with  
XX inappropriate cytokine receptor protein expression. The polypeptides  
XX may also be used as antigens in the production of antibodies, and in  
XX assays to identify modulators and inhibitors. Disorders that may be  
XX prevented, diagnosed and/or treated by the above methods include, for  
XX example immune system disorders, inflammatory disorders, immuno  
XX modulation, immune cell stimulation and proliferation, allergic  
XX reactions, aberrant cytokine receptor signalling and/or over expression  
XX of cytokine responses, cancers, cardiovascular disorders, angiogenic  
XX disorders, wound healing and regeneration, infectious diseases and  
XX neural disorders.

XX Sequence 3617 BP: 815 A: 1182 C: 1016 G: 604 T: 0 other:

Query Match 2.0%: Score 91.4; DB 22; Length 3617;  
Best Local Similarity 45.9%: Pred. No. 1.1e-15;  
Matches 388; Conservative 0; Mismatches 451; Indels 6; Gaps 2;

OY 2852 acccttctccactacacatccgcatgtatgcacaaacggatgtggcaagcagc 2911  
|||||  
Db 906 agcctctcagcttccaaatccgagtgaaagcgacaaatgacttggcagcagcag 965  
OY 2912 ccaagcaagc---gctctctacacgagcagcagcagcagcagcagcagc 2968  
|||||  
Db 966 tcaagcagagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1025  
OY 2969 aacttctctgagcagcagcagcagcagcagcagcagcagcagcagcagc 3028  
|||||  
Db 1026 tctctctctgagcagcagcagcagcagcagcagcagcagcagcagcagc 1085  
OY 3029 aacatcttcaaatatgagcagcagcagcagcagcagcagcagcagcagc 3088  
|||||  
Db 1086 agcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1145  
OY 3089 ggggtaacttccaaatccaaatccaaatccaaatccaaatccaaatccaaat 3148  
|||||  
Db 1146 aaggaactgagcagcagcagcagcagcagcagcagcagcagcagcagcag 1204  
OY 3149 cctctctctgagcagcagcagcagcagcagcagcagcagcagcagcagc 3208  
|||||  
Db 1205 -ctctctctgagcagcagcagcagcagcagcagcagcagcagcagcagc 1262  
OY 3209 cctctctctgagcagcagcagcagcagcagcagcagcagcagcagcagc 3268  
|||||  
Db 1263 tgggttct 1322  
OY 3269 accctctctgagcagcagcagcagcagcagcagcagcagcagcagcagc 3328  
|||||  
Db 1323 cagtaacttctgagcagcagcagcagcagcagcagcagcagcagcagcagc 1382  
OY 3329 ccaacttctgagcagcagcagcagcagcagcagcagcagcagcagcagc 3388  
|||||  
Db 1383 agcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1442  
OY 3389 ccaacttctgagcagcagcagcagcagcagcagcagcagcagcagcagc 3448  
|||||  
Db 1443 aagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1502  
OY 3449 tggagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3508  
|||||  
Db 1503 tgaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1562  
OY 3509 ggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3568  
|||||  
Db 1563 cgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1622  
OY 3569 gctctctctgagcagcagcagcagcagcagcagcagcagcagcagcagc 3628  
|||||

CC	DB	1623	gcgcctccagctcgtgcgaagtcagtcagctcagccacaacccctcagtcagtcgctccg	1682
OY	3629	ttccccccttcacagctgcgaagcagcatccgaagctacgttattctgctccacccct	3688	
DB	1663	aagcccgacagtlccccaatgcgactcctgagggctacagcgtggtlacagaccclgca	1742	
OY	3689	atccc 3693		
DB	1743	gcccc 1747		
RESULT 11				
ID	AAD05447/c			
XX	AAD05447	standard; CDNA: 4336 BP.		
XX	AAD05447:			
XX	17-JUL-2001	(first entry)		
XX	Human	secreted protein-encoding gene 17 cDN/ clone HHSFB67, SEQ ID NO:69		
XX	Human;	secreted protein; proliferative disorder; cancer; tumor;		
KW	foetal abnormality; developmental abnormality; haematopoietic disorder;			
KW	immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;			
KW	inflammation; allergy; neurological disorder; Alzheimer's disease;			
KW	Parkinson's disease; cognitive disorder; schizophrenia; asthma;			
KW	skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;			
KW	cardiovascular disorder; angio-genic disorder; kidney disorder;			
KW	gastrointestinal disorder; pregnancy-related disorder;			
KW	endocrine disorder; infection; wound healing; vulnerability;			
KW	cell culture; chemotaxis; food additive; gene therapy;			
XX	binding partner identification; ss.			
XX	Homo sapiens.			
PH	Key	location/Qualifiers		
FT	CDS	1297..1416		
FT		/*tag- a		
FT		/product- "Human secreted protein"		
FT		/note- "CDS does not include start codon"		
FT		/partial		
FT	sig_peptide	1297..1387		
FT		/*tag- b		
FT	mat_peptide	1388..1413		
FT		/*tag- c		
FT		/product- "Mature human secreted protein"		
XX	WO200134623-A1.			
XX	17-MAY-2001.			
XX	01-NOV-2000;	2000WO-US30037.		
XX	05-NOV-1999;	99US-0163577.		
XX	30-JUN-2000;	2000US-0215137.		
XX	(HUMA-)	HUMAN GENOME SCI INC.		
XX	Ruben SM,	Komatsoulis GA, Moore PA;		
XX	PI			
XX	WPI: 2001-316490/33.			
XX	P-PSDB: AAE01604.			
XX	Nucleic acids encoding 29 human secreted polypeptides, useful for			
XX	preventing, diagnosing and/or treating e.g. cancers, Parkinson's			
XX	disease and diabetic retinopathy -			
XX	Claim 1: Page 457-459; 535pp; English.			
XX	AAD05389-AAD05473 represent cDNAs corresponding to 29 human secreted			
XX	protein genes, and AAE01546-AAE01630 represent the proteins they encode.			
XX	AAE01631-AAE01660 represent human secreted protein fragments or variants			
XX	The secreted proteins and their genes are useful for preventing,			

treating or ameliorating medical conditions, e.g., by protein or gene therapy. Pathological conditions can be diagnosed by determining the amount of the new protein in a sample or by determining the presence of mutations in the new genes. Specific uses are described for each of the 29 genes, based on the tissues in which they are most highly expressed, and include developing products for the diagnosis and treatment of proliferative disorders, cancer, tumours, foetal and developmental abnormalities, haematopoietic disorders, diseases of the immune system, AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation, allergies, neurological disorders (e.g., Alzheimer's disease, Parkinson's disease), cognitive disorders, schizophrenia, asthma, skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders, angiogenic disorders, kidney disorders, gastrointestinal disorders, pregnancy-related disorders, endocrine disorders, and infections. The proteins can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues, to identify their cognate ligands or binding partners, and in chemotaxis, and can be used as a food additive or preservative to modify storage properties. Antibodies specific for a protein of the invention can be used in alleviating symptoms associated with the disorders mentioned above, and in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked immunosorbent assay (ELISA). The present sequence represents a human secreted protein-encoding cDNA of the invention.

Sequence 4336 BP; 766 A; 1190 C; 1357 G; 1009 T; 14 other;

Query Match	2.08	Score 91.4	DB 22	Length 4336
Best Local Similarity	45.98	Pred. NO. 1.3e-15		
Matches 388	Conservative	0	Mismatches 451	Indels 6
				Gaps 2

QY	2852	accccttcttttcaactatgagatcgcacatgtagtccaaagaacccggaattggcagaagagcagc	2911
Db	3152	AGCCCTCTACCTCTCTTAAGTTCGGAGTAAAGCCGACCATGACATTTTGCCACAGCCAGT	3093
QY	2912	ccagagcagat---gctcaccatcacagcgagcagcagcagctcctcatatgctccacctgaag	2968
Db	3092	TCAGGGAATAGTCCGAATATGACCAACCTTCGAGGCTGCCCGCATTAAGCAACCCACCA	3033
QY	2969	aagttcacctctgagcctataatcatctcagaagcctcaggtcacatlgaaagctccccaaga	3028
Db	3032	TCCTCTCCCTGATGCGCCACACACCACTACCTCTGCTATTCGATGCGAGCGCCGACACAG	2973
QY	3029	acacatcttgcacatctggaattatctgctgctatctaaatgtagtcagaagatcagcagctg	3088
Db	2972	AGGAAATATAATATATCTATCTTCTGAGATCTTCTGATATCCGGAGCTTCCTCTATCG	2913
QY	3089	ggaggttaactctccatctcaatctatctcaggtctcacaccagcgaggacagtgagtttaca	3148
Db	2912	AAGGAAATGAGGCTGATTTCAAGCTTTTCGAGGCACTGCAACACCCAGGCGCAATGGGCTGAG	2852
QY	3149	cctctgacacaccttgaaataatctcaactcagtagcagcttgtagtcagcctgtlaacggg	3208
Db	2853	--CTTACTACTCTGACACACAGCTACAGCGCTGACAGATACGATGAGCCGTATCAACGCTG	2796
QY	3209	ccgagcagcgagctctctctcttaggaataatcaccacacactctcgagaatgtgcccagtt	3268
Db	2795	TGGGTGAGGGGCGCTCCAGCCGCCCCGAGAGAGCTTTTGTGGGAGGAGCATGCCACAG	2736
QY	3269	acccccccgaagaatgtctcgaagcctagacaacatcacagaagaagcatatcaatatctggt	3328
Db	2735	CAGACACTGTGTAAGTGTGTCTCCAGCGGCCACAGGCCACACAGCTGAGACGTTCGTGG	2676
QY	3329	ccacacatttccaaggaagcctltgaatgaaatctcgaaggttccagagtcattacgtgg	3388
Db	2675	AGCCACCTCCGCTGGACAGCCAGATGAGACATCCAGGGGCTACAGATTATTTCTGGG	2616
QY	3389	ccaactctatgaaaggaagccttggttgagattaaacaatcacaccacaagcctctaac	3448
Db	2615	AACCTTACGCGGGGAGACTTCACAGCGAGTGAAGAGCGCTTTTCTCTGCTGTGAACACAGC	2556









GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 4, 2002, 00:29:06 ; Search time 226.99 Seconds  
(without alignments)  
4672.437 Million cell updates/sec

Title: US-08-956-991-1\_COPY\_453\_5135  
Perfect score: 4683  
Sequence: 1 atgtgatactgctctc.....gctgcgaggaagcagcc 4683

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 11323899 residues

al number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA: \*  
1: /cgn2\_6/ptodata/2/1na/5A.COMB.seq: \*  
2: /cgn2\_6/ptodata/2/1na/5B.COMB.seq: \*  
3: /cgn2\_6/ptodata/2/1na/6A.COMB.seq: \*  
4: /cgn2\_6/ptodata/2/1na/6B.COMB.seq: \*  
5: /cgn2\_6/ptodata/2/1na/PCTUS.COMB.seq: \*  
6: /cgn2\_6/ptodata/2/1na/backfiles1.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	529.2	11.3	1493	2 US-08-752-307B-6	Sequence 6, Appl1
2	66.4	1.4	6000	1 US-08-348-006B-6	Sequence 6, Appl1
3	66.4	1.4	6000	2 US-08-800-825A-6	Sequence 6, Appl1
4	66.4	1.4	6000	4 US-09-158-657-6	Sequence 6, Appl1
5	66.4	1.4	6000	5 PCT-US94-10166-6	Sequence 6, Appl1
6	47.8	1.0	2600	2 US-08-427-497E-4	Sequence 4, Appl1
7	47.8	1.0	3189	2 US-08-427-497E-3	Sequence 4, Appl1
8	47.8	1.0	3774	2 US-08-341-843B-1	Sequence 1, Appl1
9	47.8	1.0	3774	2 US-08-427-497E-1	Sequence 1, Appl1
10	47.8	1.0	3774	2 US-08-427-497E-2	Sequence 2, Appl1
11	45.8	1.0	5690	2 US-08-447-464-2	Sequence 2, Appl1
12	45.8	1.0	5690	2 US-08-716-679-2	Sequence 2, Appl1
13	45.2	1.0	4608	4 US-09-041-866-24	Sequence 24, Appl1
14	45.2	1.0	4608	5 PCT-US94-05277-1	Sequence 1, Appl1
15	44.6	1.0	4843	3 US-08-986-485-1	Sequence 1, Appl1
16	42.8	0.9	4975	2 US-08-249-687C-1	Sequence 1, Appl1
17	42.8	0.9	4989	2 US-08-666-392A-1	Sequence 3, Appl1
18	42.8	0.9	4989	2 US-08-625-819-1	Sequence 1, Appl1
19	42.8	0.9	4989	2 US-08-755-558-4	Sequence 4, Appl1
20	42.8	0.9	4989	3 US-08-746-559A-1	Sequence 1, Appl1
21	42.8	0.9	4989	3 US-08-880-313A-9	Sequence 9, Appl1
22	42.8	0.9	4989	4 US-09-199-926-3	Sequence 3, Appl1
23	42.8	0.9	4993	3 US-08-746-559A-3	Sequence 3, Appl1
24	42.2	0.9	3973	1 US-08-036-210-21	Sequence 21, Appl1
25	42.2	0.9	3973	2 US-08-449-609-21	Sequence 21, Appl1
26	41	0.9	1581	5 PCT-US95-08493-14	Sequence 14, Appl1
27	41	0.9	2580	5 PCT-US95-08493-18	Sequence 18, Appl1

## ALIGNMENTS

28	41	0.9	2604	5 PCT-US95-08493-20	Sequence 20, Appl1
29	40.8	0.9	7218	1 US-08-232-463-14	Sequence 14, Appl1
30	40.6	0.9	3360	1 US-08-408-093-5	Sequence 5, Appl1
31	40.6	0.9	3360	1 US-08-408-420A-5	Sequence 5, Appl1
32	40.6	0.9	3360	1 US-08-714-901-5	Sequence 5, Appl1
33	40.6	0.9	3360	3 US-08-040-741-5	Sequence 5, Appl1
34	39.2	0.8	2962	2 US-08-449-645A-10	Sequence 10, Appl1
35	39.2	0.8	2962	2 US-08-702-367A-10	Sequence 10, Appl1
36	39.2	0.8	2962	5 PCT-US95-04681-10	Sequence 10, Appl1
37	39	0.8	289	4 US-09-007-005-17	Sequence 17, Appl1
38	39	0.8	289	4 US-09-244-796-17	Sequence 17, Appl1
39	38	0.8	1004	4 US-09-383-586-24	Sequence 24, Appl1
40	36.8	0.8	1794	2 US-08-737-715-1	Sequence 1, Appl1
41	36.6	0.8	1794	2 US-08-427-497E-5	Sequence 5, Appl1
42	36.2	0.8	2869	1 US-08-374-834-2	Sequence 2, Appl1
43	36.2	0.8	2869	1 US-08-644-271-2	Sequence 2, Appl1
44	35	0.7	564	1 US-08-750-532-11	Sequence 11, Appl1
45	35	0.7	564	4 US-08-894-818B-14	Sequence 14, Appl1

RESULT 1  
US-08-752-307B-6  
Sequence 6, Application US/08752307B  
Patent No. 5952171  
GENERAL INFORMATION:  
APPLICANT: McCarthy, Sean A.  
APPLICANT: Gearing, David P.  
APPLICANT: Levinson, Douglas A.  
TITLE OF INVENTION: METHOD FOR IDENTIFYING GENES  
TITLE OF INVENTION: ENCODING NOVEL SECRETED OR MEMBRANE-ASSOCIATED PROTEIN  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson, P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: US  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows95  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/752,307B  
FILING DATE: 19-NOV-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Melkijohn, Ph.D., Anita L.  
REGISTRATION NUMBER: 35,283  
REFERENCE/DOCKET NUMBER: 09404/020001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-542-5070  
TELEFAX: 617-542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1493 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: Coding Sequence  
LOCATION: 99...1493  
US-08-752-307B-6

Query Match 11.3%: Score 529.2; DB 2; Length 1493;  
 Best Local Similarity 62.1%: Pred. No. 4.8e-150;  
 Matches 834; Conservative 0; Mismatches 508; Indels 0; Gaps 0;

```

OY 51 tgaagaccataccagcctctactcttctgcatcgtctcgaagagtgatgttgc 110
DB 152 TGAAGATGTTGGACACGACCTCTACTTGTAAATGACTCCTCCAGAGACGACCTTTTC 211
OY 111 cagacccagggagctgctgctcctcccgagcagggcctccctctgactctcag 170
DB 212 CAGCTCCCTGGGGGTGGTGTGCTGCTCCCGCGCGGCTCCCGACCGCGGCTTCG 271
OY 171 atggtacataccagcagggcagagatctacatgctcccgagctccgcaagtcaccc 230
DB 272 ATGAGTACTGGCCACAGGGGACGACATCTACAGCTGCGGCATCCGGACGCTCCACG 331
OY 231 caagcagcctctcccaattctccctctcctctccttcaagcttgatgacttaacatga 290
DB 332 CAACGGGACGCTGACGCTCTACCCCTTCCCTCCCTCCCTTCAATAGCTTTATCCACA 391
OY 291 taatacttaataatgacagctgtaaaatctcagggagaaatagaatcagatgtcca 350
DB 392 CAATGACTACTCTGACACCGCGGAGACAGCTGCGGCA GATCCGAGACCCCAATCCG 451
OY 351 catcaagcctgttttaagggagccctatacagctcgtgtgagagcagaaacatcag 410
DB 452 CGTCAAAAGCAGTTTTCAGGGAACCTTACACGCTCGGGTGAAGATCAAGCTCATCG 511
OY 411 aggcgaatgttcgtcttcaagtgatataccctccctcgtgtgagagcgtacatcgt 470
DB 512 TGGCAACGCTGGCGCTCTCAATGCTCATCCCTTCACTGAGGAGATATGTTAGGT 571
OY 471 cgtctcaatgagaaagacatcgtctcgtctcgtcagatgagatctcctacatc 530
DB 572 TGTATCTTGGGAAAGACACAGTCTCCATTCATCCAGAAACAGCTTTTATTAACCA 631
OY 531 caagcagcctgttatataaagatgacagatgaanaatgatatgatactacatcgt 590
DB 632 CCACGGCGGCTGTACATCTCTGACCTACAGAGAGAGAGCGCCCTCCACCTATCGCG 691
OY 692 CATCACCAAGACACAGTATAGCGGGAGACCGGGACAGCAATGGGGACGCTCTCTGT 751
OY 651 atcaagcagcagtaactcagcccatcactactgagatggtgtgagacatcagaaacat 710
DB 752 GACAGACCCCTGAGTGTGATCCCACTCTGATGCTTCCACTCCAGAACTGTG 811
OY 711 gctgtgagcagcgtgtgagctgctgtcaagcgtcagcctgtgagcagataltccg 770
DB 812 GCGCGGCGACACCGGTGAGCTGCTGACCGCTCGGCTTACCTATCCCGCCATCCG 871
OY 771 ctggtctgaagaaacatcgtccctctgaaacttcaagggaggttccagaagcgtgagcg 830
DB 872 CTGGCTTAAGAGTGGCGGCTCCCGGCTGACACCGCTGACCAACGATCATCAG 931
OY 831 gctgtcatgagaaacatcgtccctctgagactcagcagcgtatggtgtgagatgtcca 890
DB 932 GCTACCATGACGAGCTTGGGAGCCGAGGACGAGCGGACACTTATTTGAGGTACCAA 991
OY 891 cagatagcagacgtctagatgtagggcgcgtctgacgtgaaacacacactgaaacac 950
DB 992 CACCTTGGCTTGGCGACAGCCACAGCAATCTCATGATGATCCCTTATGATGAC 1051
OY 951 catcagtcacagaaagttaaaagcagcgtgtgtagcgaagtctctgtctcagcgt 1010
DB 1052 CTTACACCAAGAGAGTGAAGACGGCATTTGGACAGCAGGTGATCCTCTCTGCGCT 1111
OY 1011 gacagaaactgaagcagaaactctcgtgtacacgcaatgtgaaatcctcaacctg 1070
DB 1112 GACGGGCTCCCGACAGATTCCATCCCTGTGTATCCGACACGAGAGCTGTGCTGCTGA 1171

```

```

OY 1071 aaaaatgtgagagatcacagggatcaaccagaaaccttaatgagatcacatgtgtca 1130
DB 1172 CGAGGCCATCTCCATCCGAGGAGTCAAGACAGAGCTGCTCATCAGCTGGCCACAA 1231
OY 1131 aatgacggggcgcatacagatgcttctgtgcagagaaagcgtcgtcgtcaagacta 1190
DB 1232 GACGCTTCCGGGGCTTACAGTGTGCTGCTACCCGCAAGGCCAGACCGCCAGAGCTT 1291
OY 1191 tgtcagagtggtcttgaagatgaaactcccaaatattctgtgctttagtgaagagt 1250
DB 1292 TGCCATCATTTGACCTTGAGAGATGACACUCCCGCATGCTGCTCTTCAAGCAAGGT 1351
OY 1251 ggtgagtcagcagagcagcgttctcccttattgtgcaacgtgaaggaaacacttgcac 1310
DB 1352 GGTCAACCCCGGAGACACTTCTCATGATGTTGCGGCGCAAGGGCGCCCGCCAC 1411
OY 1311 gatcagtgagccctgagacatgagaccgattctcaaggtgtgagtcacgcatcagca 1370
DB 1412 GGTACCTGGGCGCTGACATGAGCCCATGCTGCGGATGAGCAGCAGCAGCAACCA 1471
OY 1371 gatgacagtcagagggagc 1392
DB 1472 GTACACCATGTGAGCAGCAGC 1493

```

## RESULT 2

```

US-08-348-006B-6
: Sequence 6, Application US/08348006B
: Patent No. 5658756
:
: GENERAL INFORMATION:
: APPLICANT: RODAN, GIDEON A.
: APPLICANT: SCHMIDT, AZRIEL
: APPLICANT: RUTLEDGE, SU JANE
: TITLE OF INVENTION: CDNA ENCODING A NOVEL HUMAN PROTEIN
: TITLE OF INVENTION: TYROSINE PHOSPHATASE
: NUMBER OF SEQUENCES: 7
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: J. MARK HAND
: STREET: 126 E. LINCOLN AVE., P.O. BOX 2000
: CITY: RAHWAY
: STATE: NEW JERSEY
: COUNTRY: USA
: ZIP: 07065-0900
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/348,006B
: FILING DATE:
: CLASSIFICATION: 514
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/122,032
: FILING DATE: 14-SEP-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: HAND, J., MARK
: REGISTRATION NUMBER: 36,545
: REFERENCE/DOCKET NUMBER: 189921A
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 908-594-3905
: TELEFAX: 908-594-4720
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 6000 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
:
: US-08-348-006B-6

```

Query Match 1.4%: Score 66.4; DB 1; Length 6000;

Best Local Similarity 47.7%; Pred. No. 2.6e-09;  
Matches 227: Conservative 0; Mismatches 246; Indels 3; Gaps 1;

```
OY 3145 tacacccctgacacactgaataagttcaactcagctgagctgagctgtaac 3204
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1721 TACGTGTGTGAGACCTGAAGCCCAACACAGGAGTACGCTTCGCGCGGCGCGTGC 1780

OY 3205 cgggcgcgcacgagggcctctctcagaagaatcaaccacacatctgagatgagcc 3264
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1781 CCGCAGGCGCTGGGCGCGCTTCACCCCGTGTGGCGGACGCGCTGCTCAACCCG 1840

OY 3265 agttacccccccgaataatgtccaaagcatagcaacatcac... aagcatatcaatcc 3324
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1841 TCAGCCCCCTCAAGACGTTAAATGTGTACGCTGCGCTCCACGCGCATTTTGTGAAGT 1900

OY 3325 tggtcacacacttccaaagaagccttgatgaattctcgaagggttcagagtcattac 3384
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1901 TGGCGCGCGCGCGCGCGGAAACGACACAGGCGCGCTGTGTGGCTACGCTCCGCTAC 1960

OY 3385 tg---gaccaactcattgagcagagagctgggtgagatttaaaaaatcacacacacag 3441
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1961 CGACCGCTGGGCTCAAGAGACCCCGAACCCCAAGAGGTGAAGCGCATCCCGACACACC 2020

OY 3442 ccttcactgagctgagcagcgtggaagaatgacacacacacacacagctgctgagc 3501
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2021 ACTCAGATCTCTGCTGGAGGCTTGGAGAGTGAACCCAGTACCGCATCAGAGTGTGCT 2080

OY 3502 ttcaaccgcgcagagagagcgggtcagagctgagatctcaccgcgacccaagaagat 3561
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2081 CACACAGAGTGGGACGACGCGCGCGAGAGCTGCGCTGTGTGCTGCCACCGAGAGAGAT 2140

OY 3562 gtccagatctctccgcggggtgtgaaggcagcgccctcagctcattgctt 3617
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2141 GTGCCACGCGCGCGCGCGGAGAGTGAAGCGGAGCGCTCAACGCCACGCGCAT 2196
```

## RESULT 3

US-08-800-825A-6

Sequence 6, Application US/08800825A

Patent No. 5866397

GENERAL INFORMATION:

APPLICANT: RODAN, GIDEON A.

APPLICANT: SCHMIDT, AZRIEL

APPLICANT: RUTLEDGE, SU JANE

TITLE OF INVENTION: CDNA ENCODING A NOVEL HUMAN PROTEIN

TITLE OF INVENTION: TYROSINE PHOSPHATASE

NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:

ADDRESSEE: J. MARK HAND - MERCK &amp; CO., INC.

STREET: 126 E. LINCOLN AVE., P.O. BOX 2000

CITY: RAHWAY

STATE: NEW JERSEY

COUNTRY: USA

ZIP: 07065-0900

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/800,825A

FILING DATE: 14-FEB-1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: HAND, J. MARK

REGISTRATION NUMBER: 36,545

REFERENCE/DOCKET NUMBER: 18992DA

TELECOMMUNICATION INFORMATION:

TELEPHONE: 732-594-3905

TELEFAX: 732-594-4720

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 6000 base pairs

TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-08-800-825A-6

Query Match 1.4%; Score 66.4; DB 2; Length 6000;

Best Local Similarity 47.7%; Pred. No. 2.6e-09;  
Matches 227: Conservative 0; Mismatches 246; Indels 3; Gaps 1;

```
OY 3145 tacacccctgacacactgaataagttcaactcagctgagctgagctgtaac 3204
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1721 TACGTGTGTGAGACCTGAAGCCCAACACAGGAGTACGCTTCGCGCGGCGCGTGC 1780

OY 3205 cgggcgcgcacgagggcctctctcagaagaatcaaccacacatctgagatgagcc 3264
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1781 CCGCAGGCGCTGGGCGCGCTTCACCCCGTGTGGCGGACGCGCTGCTCAACCCG 1840

OY 3265 agttacccccccgaataatgtccaaagcatagcaacatcac... aagcatatcaatcc 3324
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1841 TCAGCCCCCTCAAGACGTTAAATGTGTACGCTGCGCTCCACGCGCATTTTGTGAAGT 1900

OY 3325 tggtcacacacttccaaagaagccttgatgaattctcgaagggttcagagtcattac 3384
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1901 TGGCGCGCGCGCGCGCGGAAACGACACAGGCGCGCTGTGTGGCTACGCTCCGCTAC 1960

OY 3385 tg---gaccaactcattgagcagagagctgggtgagatttaaaaaatcacacacacag 3441
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1961 CGACCGCTGGGCTCAAGAGACCCCGAACCCCAAGAGGTGAAGCGCATCCCGACACACC 2020

OY 3442 ccttcactgagctgagcagcgtggaagaatgacacacacacacacagctgctgagc 3501
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2021 ACTCAGATCTCTGCTGGAGGCTTGGAGAGTGAACCCAGTACCGCATCAGAGTGTGCT 2080

OY 3502 ttcaaccgcgcagagagagcgggtcagagctgagatctcaccgcgacccaagaagat 3561
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2081 CACACAGAGTGGGACGACGCGCGCGAGAGCTGCGCTGTGTGCTGCCACCGAGAGAGAT 2140

OY 3562 gtccagatctctccgcggggtgtgaaggcagcgccctcagctcattgctt 3617
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2141 GTGCCACGCGCGCGCGCGGAGAGTGAAGCGGAGCGCTCAACGCCACGCGCAT 2196
```

## RESULT 4

US-09-158-657-6

Sequence 6, Application US/09158657

Patent No. 6214564

GENERAL INFORMATION:

APPLICANT: RODAN, GIDEON A.

APPLICANT: SCHMIDT, AZRIEL

APPLICANT: RUTLEDGE, SU JANE

TITLE OF INVENTION: CDNA ENCODING A NOVEL HUMAN PROTEIN

TITLE OF INVENTION: TYROSINE PHOSPHATASE

NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:

ADDRESSEE: J. MARK HAND - MERCK &amp; CO., INC.

STREET: 126 E. LINCOLN AVE., P.O. BOX 2000

CITY: RAHWAY

STATE: NEW JERSEY

COUNTRY: USA

ZIP: 07065-0900

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/158,657

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/800,825

1 FILING DATE: 14-FEB-1997  
 2 ATTORNEY/AGENT INFORMATION:  
 3 NAME: HAND, J. MARK  
 4 REGISTRATION NUMBER: 36,545  
 5 REFERENCE/DOCKET NUMBER: 18992DA  
 6 TELECOMMUNICATION INFORMATION:  
 7 TELEPHONE: 732-594-3905  
 8 TELEFAX: 732-594-4720  
 9 INFORMATION FOR SEQ ID NO: 6:  
 10 SEQUENCE CHARACTERISTICS:  
 11 LENGTH: 6000 base pairs  
 12 TYPE: nucleic acid  
 13 STRANDEDNESS: single  
 14 TOPOLOGY: linear  
 15 MOLECULE TYPE: cDNA  
 16 US-09-158-657-6

Query Match	1.48;	Score 66.4;	DB 4;	Length 6000;
1st Local Similarity	47.7%;	Pred. No. 2.6e-09;		
atches 227; Conservative	0;	Mismatches 246;	Indels 3;	Gaps 1

OY	3145	tacacccctgacacacttgataatgaatcactcagtaagcgcgtggtggtgtccagagctgtgaac	3204
Db	1721	TACCTGTGTGAGAGACCTTGAAGCCCAACACGGAGTAgGCC'TTCGCTTGGCGGCCGCTGC	1780
OY	3205	cgggcgcgcacagggagcctctctctcgaagaatcattcaccacactctcagagatgtagcc	3264
Db	1781	CCGAGAGGCGCTGGCGCGCTTCAACCCCGTGGTGGCGACACCGCAGCTGCTACGTCGAACCG	1840
OY	3265	agttacccccccggaanaatgtccaaagccatagcaacataccagcaagaagatctcaatattcc	3324
Db	1841	TCAGACCCCCCTTCAMAGACGTTAAATGTGTAGCGTGGCTCCACGGCCATTTTGTGTAAGT	1900
OY	3325	tgtgtccacactcttccaaagagcctctgaaatgtgaattctccaggggtltcaagatcaltac	3384
Db	1901	TGGCGCCCGCGCGCGCGGAAACGCAACAGGGGCCCTGTGTGGGTACAGCGTCCGGCTAC	1960
OY	3385	tg---ggccaacactcatgtagcagagagcctggtgtgagat aaaatattcaaccaccacag	3441
Db	1961	CGACCGCTGGGCTCTAGAGAGCCCGGGAACCCAMAGAGTATGACAGCGATCCCCCGACACAC	2020
OY	3442	ccttcaactgagactgtgagacagcctgtgaaanaatcaccaactacacagatccaggtgtctggcc	3501
Db	2021	ACTCAGATCTCTGCTGGAGAGCCTTGGAAATGTGAAACCATACCGGATCAGCACTGTCCGT	2080
OY	3502	ttcaccccgctgacagagacaggtgtcagagatgtgacagatcttcaaccggacccaagaagat	3561
	2081	CACACAGAGAGTGGGACACGAGCGCCCGAGAGCTTCCCGGTGTCTCTCCGACCCGACGAGAT	2140
OY	3562	gttccaggtctctccgcgggtgtgtaaagcaagcgagcgccctcagcctcgaatgtctt	3617
Db	2141	GTGCCACACGCGCGCGCCCGGAAGGTGAGACCGGAGGCGCTCTCAACGCCACACGGCCAT	2196

```

1  RESULT 5
2  PCT-US94-10166-6
3  : Sequence 6, Application PC/TUS9410166
4  :
5  : GENERAL INFORMATION:
6  :
7  : APPLICANT: RODAN, GIDEON A
8  : APPLICANT: SCHMIDT, AZRIEL
9  : APPLICANT: RUTLEDGE, SU JANE
10 :
11 : TITLE OF INVENTION: CDNA ENCODING A NOVEL HUMAN PROTEIN
12 :
13 : TITLE OF INVENTION: TYROSINE PHOSPHATASE
14 :
15 : NUMBER OF SEQUENCES: 6
16 :
17 : CORRESPONDENCE ADDRESSES:
18 :
19 : ADDRESSEE: JOHN W. WALLEN III
20 :
21 : STREET: P.O. BOX 2000, 126 E. LINCOLN AVE.
22 :
23 : CITY: RAHWAY
24 :
25 : STATE: NJ
26 :
27 : COUNTRY: USA
28 :
29 : ZIP: 07065
30 :
31 : COMPUTER READABLE FORM:

```

```

MEDIC TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/10166
FILING DATE: 09-SEPT-1994
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/122,032
FILING DATE: 14-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: WALLEN, JOHN W III
REGISTRATION NUMBER: 35403
REFERENCE/DOCKET NUMBER: 18992
TELECOMMUNICATION INFORMATION:
TELEPHONE: 908-594-3905
TELEFAX: 908-594-4720
TELEX: 138825
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 6000 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
PCT-US94-10166-6

```

Query Match	1.4%	Score 66.4	DB 5	Length 6000
Best Local Similarity	47.7%	Pred No. 2.6e-09		
Matches 227	Conservative	0	Mismatches 246	Indels 3
				Gaps 1

QY	3145	tacacccctggagcaaaccttggaaatgaattctacatcaga taagcgctctgg tggctcaaggccctgaac	3204
Db	1721	TACGTGTGGAGAGACCTGGAAGCCCAACAGCAATACGCTTTGGCTTGGCGGCGCGCTCG	1780
QY	3205	cgggcgcgcagcaggagcctctctctcagaaatcatcaccaactctcagagatgtgtccc	3264
Db	1781	CCGAGCGCGCTGTGGCGGCTTCACCCCTGTGTGTGGCGACGCGACGCGCTGTGAGCCAAACCG	1840
QY	3265	agtaacccccccgaataatgttccaaagcatatgcaacatcaacgaaagacatcatcaatctcc	3324
Db	1841	TCAGGCCCCCTCCAAAGACGTTAAATGTGTAGCGTGGCCTCCACGGCCATTTTGGTAAGT	1900
QY	3325	tgtctcacacatttccaaagagccttgaatgaattctccagaggtctcagagtcattatc	3384
Db	1901	TGGCGGCTGTGCCGCTCCCGCGAAMGCCAACAGGGGGCCCTGGTGGGTACAGCTCCGCTAC	1960
QY	3385	tg---gggtcaatctcatatgaatctgaatcttggatctggatctaaataacttcaacacacagag	3444
Db	1961	CCAGCCCTCTGGCTTCAGAGGATCTGGGAATCCACAGAGAGTGAACGGCTATCCCCCGACACAC	2020
QY	3442	cgttcaacttggaaactctgagacaggtcttggaaagtatcaccaactacacagatctcagg tctctgccc	3501
Db	2021	ACTCAAGTCTCTGCTGGAGGCTTTGGAGAAGTGGACCCAGTACCGCATCAAGACTGTGCGT	2080
QY	3502	ttcaccttgcagcagagagacagagatctcagagatgtagcagaatcttcaacccggacacaaagagat	3561
Db	2081	CACACAGAGAGTGGGACCAAGGCTCCGAGAGCTCCGCCCTGTGTCTCCGACACCGACGAGGAT	2140
QY	3562	gttccagagctctccctccgggggtcttgaagagcagctgagcctcagaactctcagatgtctt	3617
Db	2141	GTGCCCCAGTGCCGCCCGCCGAGAGCTGAGAGCGGAGGCGCTCAACGCCACAGGCCAT	2196

RESULT 6  
US-08-427-497E-4  
; Sequence 4, Application US/08427497E  
; Patent No. 5968124  
; GENERAL INFORMATION:  
; APPLICANT: Iemmon, Vance  
; TITLE OF INVENTION: A Method for Characterizing the

TITLE OF INVENTION: Nucleotide Sequence of L1CAM and  
 Patent No. 5969124  
 TITLE OF INVENTION: the Nucleotide Sequence  
 TITLE OF INVENTION: Characterized Thereby  
 NUMBER OF SEQUENCES: 44  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Fay, Sharpe, Beall, Fagan,  
 ADDRESS: Minnich & McKee  
 STREET: 1100 Superior Avenue  
 STREET: Suite 700  
 CITY: Cleveland  
 STATE: Ohio  
 COUNTRY: U.S.A.  
 ZIP: 44114-2518  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb  
 MEDIUM TYPE: storage  
 COMPUTER: Compaq Prolinea 5100e  
 OPERATING SYSTEM: DOS 5.0  
 SOFTWARE: ASCII  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/427,497E  
 FILING DATE: April 24, 1995  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 07/904,991  
 FILING DATE: June 26, 1992  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Minnich, Richard J.  
 REGISTRATION NUMBER: 24,175  
 REFERENCE/DOCKET NUMBER: CWR 2 149-3-1  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (216) 861-5582  
 TELEFAX: (216) 241-1666  
 TELEX: (216) 980162  
 INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2600  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: nucleic acids  
 HYPOTHETICAL: irrelevant  
 ANTI-SENSE: no  
 ORIGINAL SOURCE:  
 ORGANISM: homo sapiens  
 INDIVIDUAL ISOLATE: 17-18 week fetus  
 IMMEDIATE SOURCE:  
 LIBRARY: Stratagene cDNA library 936206  
 CLONE: 4  
 PUBLICATION INFORMATION:  
 AUTHORS: Blavin, Mary Louise  
 AUTHORS: Lemmon, Vance  
 TITLE: Molecular structure and functional testing of  
 TITLE: human L1CAM: an interspecies comparison.  
 JOURNAL: GENOMICS  
 VOLUME: 11  
 ISSUE:  
 PAGES: 416-423  
 DATE: 1991  
 RELEVANT RESIDUES IN SEQ ID NO: 1108 to 3708

[illegible]

Db	1615	GAGTTCACTTTA---GCACCCAGAGGGAGTGCCTCGGGACCCCGAGGGCGTTTGACCTG	1671
Qy	3592	gcgcgcgcctccaaagcctccaaatcgcgcgtcttgctccgcctcccccctcaagtgaagcgc	3651
Db	1672	GAGGGCCACTGGAAACACACACCTGCTGCTGCGCTGGACAGCCCCCACTGACGCCAACAAGGC	1731
Qy	3652	atcatcgaagaatlaactgtatattctgcctccaccccat	3690
Db	1732	GTGCTCAACCGGCTACGCTGCTCTCTACCAACCCCTTGAT	1770

RESULT 7  
 US-08-427-497E-3  
 Sequence 3, Application US/08427497E  
 Patent No. 5969124  
 GENERAL INFORMATION:  
 APPLICANT: Iemmon, Vance  
 TITLE OF INVENTION: A Method for Characterizing the  
 TITLE OF INVENTION: Nucleotide Sequence of L1CAM and  
 Patent No. 5969124  
 TITLE OF INVENTION: the Nucleotide Sequence  
 TITLE OF INVENTION: Characterized thereby  
 NUMBER OF SEQUENCES: 44  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Fay, Sharpe, Beall, Pagan,  
 ADDRESSEE: Minnich & McKee  
 STREET: 1100 Superior Avenue  
 STREET: Suite 700  
 CITY: Cleveland  
 STATE: Ohio  
 COUNTRY: U.S.A.  
 ZIP: 44114-2518  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb  
 MEDIUM TYPE: storage  
 COMPUTER: Compaq Prolinea 5100e  
 OPERATING SYSTEM: DOS 5.0  
 SOFTWARE: ASCII  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/427,497E  
 FILING DATE: April 24, 1995  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 07/904,991  
 FILING DATE: June 26, 1992  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Minnich, Richard J.  
 REGISTRATION NUMBER: 24,175  
 REFERENCE/DOCKET NUMBER: CWR 2 149,3-1  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (216) 861-5582  
 TELEFAX: (216) 241-1666  
 TELEX: (216) 980162  
 INFORMATION FOR SEQ ID NO: 3:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 3189  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: nucleic acids  
 HYPOTHETICAL: irrelevant  
 ANTI-SENSE: no  
 ORIGINAL SOURCE:  
 ORGANISM: homo sapiens  
 INDIVIDUAL ISOLATE: 17-18 week fetus  
 IMMEDIATE SOURCE:  
 LIBRARY: Stratagene cDNA Library 936206  
 CLONE: 3.1  
 PUBLICATION INFORMATION:  
 AUTHORS: Hlavin, Mary Louise  
 AUTHORS: Iemmon, Vance  
 TITLE: Molecular structure and functional testing of



```

1 MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
2 MEDIUM TYPE: Storable
3 COMPUTER: Compaq Prolinea 5100e
4 OPERATING SYSTEM: DOS 5.0
5 SOFTWARE: ASCII
6 CURRENT APPLICATION DATA:
7 APPLICATION NUMBER: US/08/427,497E
8 FILING DATE: April 24, 1995
9 CLASSIFICATION: 435
10 PRIOR APPLICATION DATA:
11 APPLICATION NUMBER: 07/904,991
12 FILING DATE: June 26, 1992
13 ATTORNEY/AGENT INFORMATION:
14 NAME: Minnich, Richard J.
15 REGISTRATION NUMBER: 24,175
16 REFERENCE/DOCKET NUMBER: CMR 2 149-3-1
17 TELECOMMUNICATION INFORMATION:
18 TELEPHONE: (216) 861-5582
19 TELEFAX: (216) 241-1666
20 TELEX: (216) 980162
21 INFORMATION FOR SEQ ID NO: 1:
22 SEQUENCE CHARACTERISTICS:
23 LENGTH: 3774
24 TYPE: nucleic acid
25 STRANDEDNESS: single
26 TOPOLOGY: linear
27 MOLECULE TYPE: nucleic acids
28 HYPOTHETICAL: irrelevant
29 ANTI-SENSE: no
30 ORIGINAL SOURCE:
31 ORGANISM: Homo Sapiens
32 INDIVIDUAL ISOLATE: 17-18 week fetus
33 IMMEDIATE SOURCE:
34 LIBRARY: Stratagene cDNA Library 936206
35 CLONE: synthesis of 4 clones
36 PUBLICATION INFORMATION:
37 AUTHORS: Hlavin, Mary Louise
38 AUTHORS: Lemmon, Vance
39 TITLE: Molecular structure and functional
40 TITLE: testing of human L1CAM: an
41 TITLE: interspecies comparison.
42 JOURNAL: GENOMICS
43 VOLUME: 11
44 ISSUE:
45 PAGES: 416-423
46 DATE: 1991
47 RELEVANT RESIDUES IN SEQ ID NO: 1 to 3774
48
49 US-08-427-497E-1

```

Db 2839 GTGCTACCGGCTACGTGCTCTCTCTACACCCCTGGAT 2877

Sequence 2, Application US/08427497E  
Patent No. 5969124  
GENERAL INFORMATION:  
APPLICANT: Lemmon, Vance  
TITLE OF INVENTION: A Method for Characterizing the  
TITLE OF INVENTION: Nucleotide Sequence of L1CAM and  
Patent No. 5969124  
TITLE OF INVENTION: the Nucleotide Sequence  
TITLE OF INVENTION: Characterized Thereby  
NUMBER OF SEQUENCES: 44  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fay, Sharpe, Beall, Fagan,  
ADDRESSEE: Munich 8 McKee  
STREET: 1100 Superior Avenue  
STREET: Suite 700  
CITY: Cleveland  
STATE: Ohio  
COUNTRY: U.S.A.  
ZIP: 44114-2518  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb  
MEDIUM TYPE: Storable  
COMPUTER: Compaq Prolinea 5100e  
OPERATING SYSTEM: DOS 5.0  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/427,497E  
FILING DATE: April 24, 1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/904,991  
FILING DATE: June 26, 1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Munich, Richard J.  
REGISTRATION NUMBER: 24,175  
REFERENCE/DOCKET NUMBER: CWR 2 149-3-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (216) 861-5582  
TELEFAX: (216) 241-1666  
TELEX: (216) 980162  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERIST: S:  
LENGTH: 3774  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: nucleic acids  
HYPOTHETICAL: irrelevant  
ANTI-SENSE: no  
ORIGINAL SOURCE:  
ORGANISM: Homo Sapiens  
INDIVIDUAL ISOLATE: T7-18 week fetus  
IMMEDIATE SOURCE:  
LIBRARY: Stratagene cDNA Library 936206  
CLONE: synthesis of 4 clones  
PUBLICATION INFORMATION:  
AUTHORS: Hlavio, Mary Louise  
AUTHORS: Lemmon, Vance  
TITLE: Molecular structure and functional testing of  
TITLE: human L1CAM: an interspecies comparison.  
JOURNAL: GENOMICS  
VOLUME: 11  
ISSUE:  
PAGES: 416-423  
DATE: 1991  
RELEVANT RESIDUES IN SEQ ID NO: 1 to 3774  
US-08-427-497E-2

Query Match	1.0%;	Score 47.8;	DB 2;	Length 3774;
Best Local Similarity	54.3%;	Pred. NO. 0.00082;		
Matches 119;	Conservative 0;	Mismatches 97;	Indels 3;	Gaps 1;



QY	3532	gagcaga	tctctacccgagacaaagagat	tgltc	aagtc	ctctcccg	gagtg	gtgaagca	3531
QY <td>2722</td> <td>GAGTTCACCTTGA</td> <td>---GAC</td> <td>CCGAGGGA</td> <td>GTGGCTGG</td> <td>CCACCC</td> <td>CCCGGCTTGG</td> <td>ATGACCTG</td> <td>2778</td>	2722	GAGTTCACCTTGA	---GAC	CCGAGGGA	GTGGCTGG	CCACCC	CCCGGCTTGG	ATGACCTG	2778
DB <td>2662</td> <td>TATAGCTTCCTAC</td> <td>CACCTG</td> <td>AGAGGTG</td> <td>CAGGCGCTTTA</td> <td>AGGGCG</td> <td>GAGGATCG</td> <td>GGGCGCCG</td> <td>CAGC 2721</td>	2662	TATAGCTTCCTAC	CACCTG	AGAGGTG	CAGGCGCTTTA	AGGGCG	GAGGATCG	GGGCGCCG	CAGC 2721
QY <td>3472</td> <td>taccacaa</td> <td>taagatctcaag</td> <td>gtgctg</td> <td>agcttacc</td> <td>ccgacg</td> <td>cgaggaag</td> <td>aaagtg</td> <td>taagagtl 3531</td>	3472	taccacaa	taagatctcaag	gtgctg	agcttacc	ccgacg	cgaggaag	aaagtg	taagagtl 3531
QY <td>3652</td> <td>atcalccga</td> <td>aaftacac</td> <td>tgatctg</td> <td>ctccac</td> <td>accctt</td> <td></td> <td></td> <td>3650</td>	3652	atcalccga	aaftacac	tgatctg	ctccac	accctt			3650
DB <td>2839</td> <td>GTGCTCACCGCT</td> <td>ACAGCGCTCT</td> <td>CTAC</td> <td>CAACCC</td> <td>CCCTGGAT</td> <td></td> <td></td> <td>2877</td>	2839	GTGCTCACCGCT	ACAGCGCTCT	CTAC	CAACCC	CCCTGGAT			2877

RESULT 11  
US-08-447-464-2  
Sequence 2, Application US/08447464  
Patent No. 5840842  
GENERAL INFORMATION:

```

1 GENERAL INFORMATION:
2 APPLICANT: Schlessinger, Joseph
3 APPLICANT: Yan, Hai
4 TITLE OF INVENTION: NOVEL RECEPTOR-TYPE PROTEIN
5 TITLE OF INVENTION: PHOSPHOTYROSINE PHOSPHATASE-SIGMA
6 NUMBER OF SEQUENCES: 12
7 CORRESPONDENCE ADDRESS:
8 ADDRESSEE: Pennie & Edmonds
9 STREET: 1155 Avenue of the Americas
10 CITY: New York
11 STATE: New York
12 COUNTRY: U.S.A.
13 ZIP: 10036-2711
14
15 COMPUTER READABLE FORM:
16 MEDIUM TYPE: FLOPPY disk
17 COMPUTER: IBM PC compatible
18 OPERATING SYSTEM: PC-DOS/MS-DOS
19 SOFTWARE: PatentIn Release #1.0, Version #1.25
20 CURRENT APPLICATION DATA:
21 APPLICATION NUMBER: US/08/447,464
22 FILING DATE: 24-MAY-1995
23 CLASSIFICATION: 435
24
25 PRIOR APPLICATION DATA:
26 APPLICATION NUMBER: 08/130,570
27 FILING DATE: 01-OCT-1993
28 ATTORNEY/AGENT INFORMATION:
29 NAME: Mistrock, S. Leslie
30 REGISTRATION NUMBER: 18,872
31 REFERENCE/DOCKET NUMBER: 7683-043
32
33 TELECOMMUNICATION INFORMATION:
34 TELEPHONE: 212-790-9090
35 TELEFAX: 212-869-8864/9741
36
37 TELEX: 66141 PENNIE
38 INFORMATION FOR SEQ ID NO: 2:
39 SEQUENCE CHARACTERISTICS:
40 LENGTH: 5690 base pairs
41 TYPE: nucleic acid
42 STRANDEDNESS: unknown
43 TOPOLOGY: unknown
44 MOLECULE TYPE: DNA (genomic)
45 FEATURE:
46 NAME/KEY: CDS
47 LOCATION: 833..5338
48
49 US-08-447-464-2

```

```

Query Match      1 0%; Score 45.8; DB 2; Length 5690;
Best Local Similarity 51.7%; Pred. No. 0.0044;
Matches 104; Conservative 0; Mismatches 97; Indels 0; Gaps 0.
0Y 3455 tggacgagcctgtaaaaglacaccactatcacatccatgttgcctgcttccacgcgcag 3514
||||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Db	2236	TGGCACCCTCGTAGAGAGATGAGACCTAACACTGTGAGAGTGTCTGGCTTCACATCGGTGG	2315
Qy	3515	gaacagagagatcagagatagacagatcttccacccggaaccaagaagatgttccaggtcttc	3574
Db	2316	GGCATGGGACCACTGTGACACCCCATCAGTCCAGTCCAMGACCCAGAGGAGTGGCCCGCCAGC	2375
Qy	3575	ccgagagatgtgaagagcagcagcgagccctcagagctccatggtcttctgtctcgtgctccccc	3634
Db	2376	GCATGAAATTTCGGCTGAGAGCCACAGTCAGAGACCCAGCATGTGGCTCTCTGTGAGTGCAC	2435
Qy	3635	ctcttcaagcttgaacgagatca	3655
Db	2436	CACCGCTACGAGAGATGTCTCATTA	2456

RESULT 12  
US-08-716-679-2  
; Sequence 2, Application 1 /08716679

```

1 Patent NO. 5846800
2 GENERAL INFORMATION:
3 APPLICANT: Schlessinger, Joseph
4 APPLICANT: Yan, Hai
5 TITLE OF INVENTION: NOVEL RECEPTOR-TYPE PROTEIN
6 TITLE OF INVENTION: PHOSPHOTRANSFERASE PHOSPHATASE-SIGMA
7 NUMBER OF SEQUENCES: 12
8 CORRESPONDENCE ADDRESS:
9 ADDRESSEE: Pennie & Edwards
10 STREET: 1155 Avenue of the Americas
11 CITY: New York
12 STATE: New York
13 COUNTRY: U.S.A.
14 ZIP: 10036-2711
15 COMPUTER READABLE FORM:
16 MEDIUM TYPE: Floppy disk
17 COMPUTER: IBM PC compatible
18 OPERATING SYSTEM: PC-DOS/MS-DOS
19 SOFTWARE: PatentIn Release #1.0, Version #1.25
20 CURRENT APPLICATION DATA:
21 APPLICATION NUMBER: US/08/716,679
22 FILING DATE:
23 CLASSIFICATION: 435
24 PRIOR APPLICATION DATA:
25 APPLICATION NUMBER: US/08/130,570
26 FILING DATE:
27 ATTORNEY/AGENT INFORMATION:
28 NAME: Mistrock, S. Jessie
29 REGISTRATION NUMBER: 18,872
30 REFERENCE/INCKET NUMBER: 7684,044
31 TELECOMMUNICATION INFORMATION:
32 TELEPHONE: 212-790-9090
33 TELEFAX: 212-869-8864/9141
34 TELEX: 06141 PENNIE
35 INFORMATION FOR SEQ ID NO: 2:
36 SEQUENCE CHARACTERISTICS:
37 LENGTH: 5690 base pairs
38 TYPE: nucleic acid
39 STRANDEDNESS: unknown
40 TOPOLOGY: unknown
41 MOLECULE TYPE: DNA (genomic)
42 FEATURE:
43 NAME/KEY: CDS
44 LOCATION: 833..5338
45 US-08-716-679-2

```

Query Match	1.0%	Score 45.8	DB 2	Length 5690
Best Local Similarity	51.7%	Pred No. 0.0044		
Matches 104	Conservative 0	Mismatches 97	Indels 0	Gaps 0

  

QY	3455	lygaacgqctgtaagaagylacacccaactacacgacalccagctgctgcgctcttaccgcggag	3514
Db	2256	TGGGCACTCTTCTTAGAGAGATGAAGACCTCACTCTGTGAAGAGTCTCCGCTTTCACATCCGTGG	2315





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 4, 2002, 02:14:32 ; Search time 6206.58 Seconds  
(without alignments)  
34.627 Million cell updates/sec

Title: US-08-956-991-5

Perfect score: 20

Sequence: 1 ccaagttcacaagagcagg 20

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues 22703874

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

EST:  
1: em\_estfun:  
2: em\_esthum:  
3: em\_estlun:  
4: em\_estlun:  
5: em\_estlun:  
6: em\_estlun:  
7: em\_estlun:  
8: em\_estlun:  
9: em\_hlc:  
10: gb\_estl:  
11: gb\_estl:  
12: gb\_hlc:  
13: gb\_gss:  
14: em\_gss\_fun:  
15: em\_gss\_hum:  
16: em\_gss\_inv:  
17: em\_gss\_pln:  
18: em\_gss\_pro:  
19: em\_gss\_rtd:  
20: em\_gss\_vrt:  
21: em\_gss\_other:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	744	11	BF348620 602019924
2	18.4	92.0	538	11	BG383716 301966 MA
3	17.4	87.0	442	11	BG348199 d939c10.Y
4	17.4	87.0	739	11	BT102963 60288573
5	17.4	87.0	832	13	CNS025H0
6	16.8	84.0	177	10	AM176762 Tetradon
7	16.8	84.0	319	10	AV158019 AV158019
8	16.8	84.0	334	10	AT422998 tf31b11.x
9	16.8	84.0	343	10	AT927023 w672a06.x
10	16.8	84.0	350	10	AM780736 s184f06.Y
11	16.8	84.0	356	11	N24917 YY01h12.S1
12	16.8	84.0	358	10	AM015308 UI-H-B10-

C 13	16.8	84.0	359	11	C84007	C84007 C84007 Dict
C 14	16.8	84.0	380	10	A1347190	A1347190 tc04g04.x
C 15	16.8	84.0	380	13	A0996872	RPCT-23-3
C 16	16.8	84.0	393	10	AA200841	mu03g06.r
C 17	16.8	84.0	400	13	A0894484	HS-2036.B
C 18	16.8	84.0	444	13	A2363648	AA188488 zq44d06.s
C 19	16.8	84.0	449	10	AA188488	AA188488 zq44d06.s
C 20	16.8	84.0	467	13	A2229092	RPCT-23-6
C 21	16.8	84.0	469	10	A1636091	ts92f09.x
C 22	16.8	84.0	475	13	A2363734	IM0109F20
C 23	16.8	84.0	502	13	A0150756	HS-3091.A
C 24	16.8	84.0	502	11	B1426424	8ag03a12.
C 25	16.8	84.0	526	10	AA160797	AA160797 z060e04.s
C 26	16.8	84.0	583	13	A0307774	HS-2175.A
C 27	16.8	84.0	603	10	AA484296	AA484296 GWT10001B
C 28	16.8	84.0	710	11	BG51248	602730380
C 29	16.8	84.0	710	13	A2717126	RPCT-24-1
C 30	16.8	84.0	725	10	AU005349	AU005349
C 31	16.8	84.0	732	11	BG480910	602530058
C 32	16.8	84.0	756	13	A272263	RPCT-23-1
C 33	16.8	84.0	762	11	B193278	602946987
C 34	16.8	84.0	804	11	B1145737	BT145737
C 35	16.8	84.0	826	13	CNS038RC	AL232877 Tetradon
C 36	16.8	84.0	826	13	AQ739885	HS-5505.A
C 37	16.8	84.0	880	11	B1412208	B1412208 602990178
C 38	16.8	84.0	938	10	BE253944	BE253944 601113242
C 39	16.8	84.0	993	13	CNS060CY	AL410504 T3 end of
C 40	16.8	84.0	1222	11	BF688244	BF688244 602184789
C 41	16.8	84.0	1224	10	AA694519	AA694519 ab35h05.s
C 42	16.4	82.0	312	13	AQ339036	HS-5021.A
C 43	16.4	82.0	374	11	BF391154	UI-R-CA1-
C 44	16.4	82.0	376	11	BI295719	BI295719 UI-R-DK0-
C 45	16.4	82.0	390	10	AA770324	AA770324 ab83a04.s

#### ALIGNMENTS

RESULT 1  
LOCUS BF348620 114 bp mRNA EST 22-NOV-2000  
DEFINITION 60201:924F1 NCI CGAP, Brn67 Homo sapiens CDNA clone IMAGE:4155508  
5', mRNA sequence.  
ACCESSION BF348620  
VERSION BF348620.1 GI:11296215  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 744)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
CONTACT: Robert Strausberg, Ph.D.  
Email: cgaps-remail.nih.gov  
Tissue Procurement: David N. Louis, M.D.  
CDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium (LLNL)  
http://image.llnl.gov  
Plate: L1AM9426 row: P column: 05  
High quality sequence stop: 697.  
Location/Qualifiers:  
1..744  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4155508"  
/clone\_lib="NCI CGAP Brn67"  
/tissue\_type="anaplastic oligodendroglioma with 1p/19q  
loss"

/lab\_host="DHI08 (T1 phage-resistant)"  
 /note="Organ: brain; Vector: PCMV-SPORT6; Site: 1; NotI;  
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
 Average insert size 2.3 kb. Constructed by Life  
 Technologies. Note: this is a NCI-CGAP Library."  
 BASE COUNT 188 a 196 c 210 g 150 t  
 ORIGIN

Query Match 100.0%; Score 20; DB 11; Length 744;  
 Best Local Similarity 100.0%; Pred. No. 23;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ccagttctcaagagcagc 20  
 |||  
 DB 20 CCAGTCTCAAGAGCAGC 39

RESULT 2  
 BG383716 538 bp mRNA EST 12-MAR-2001  
 DE-NTITION 301966 MARC 1P1G Sus scrofa cDNA 5', mRNA sequence.  
 ACCESSION BG383716  
 VERSION BG383716.1 GI:13308188  
 KEYWORDS EST.  
 SOURCE Pig.  
 ORGANISM Sus scrofa  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 1 (bases 1 to 538)  
 Fahnenkrug,S.C., Frelking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,  
 Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegfeld,W.W.,  
 and Keele,J.W.  
 Design and use of two pooled tissue normalized cDNA libraries for  
 EST discovery in swine  
 Unpublished (2000)  
 JOURNAL  
 COMMENT USDA, ARS, US Meat Animal Research Center  
 Contact: Smith TPL  
 PO Box 166, Clay Center, NE 68933-0166, USA  
 Tel: 402 762 4366  
 Fax: 402 762 4390  
 Email: smith@emall.marc.usda.gov  
 Single pass sequencing. Bases called and alt-trimmed with phred  
 v0.980904.e. Vector identified by cross-match with the -minscore 18  
 and -minmatch 12 options.  
 PCR primers  
 FORWARD: AGGAACAGCTATGACCAT  
 BACKWARD: GTTTCACAGCAGCAGC  
 Plate: 88 row: K column: 9  
 Seq primer: ATTAGTGACACTATAG.  
 Location/Qualifiers  
 1..538  
 /organism="Sus scrofa"  
 /db\_xref="taxon:9623"  
 /clone\_id="MARC 1P1G"  
 /tissue\_type="pooled"  
 /lab\_host="DHI08"  
 /note="Vector: PCMV SPORT6; Site: 1; XbaI; Site 2: XhoI;  
 library made from pooled tissue from day 11, 13, 15, 20,  
 and 30 embryos."

BASE COUNT 98 a 150 c 120 g 170 t  
 ORIGIN

Query Match 92.0%; Score 18.4; DB 11; Length 538;  
 Best Local Similarity 95.0%; Pred. No. 1.3e+02;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ccagttctcaagagcagc 20  
 |||  
 DB 178 CCAGTCTCAAGAGCAGC 197

RESULT 3  
 BG348199 442 bp mRNA EST 18-APR-2001  
 LOCUS BG348199  
 DEFINITION d939c10.y1 Xenopus laevis gastrula non normalized Xenopus laevis  
 cDNA clone IMAGE:3751290 5', mRNA sequence.  
 ACCESSION BG348199  
 VERSION BG348199.1 GI:13168625  
 KEYWORDS EST.  
 SOURCE African clawed frog.  
 ORGANISM Xenopus laevis  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;  
 Xenopodinae; Xenopus.  
 1 (bases 1 to 442)  
 Clifton,S., Johnson,S.L., Blumberg,B., Song,J., Hillier,L., Pape,D.,  
 Martin,J., Wylie,T., Underwood,K., Theising,B., Bowers,Y., Person  
 ,B., Gibbons,M., Harvey,N., Riller,E., Jackson,T., McCann,R.,  
 Waterston,R. and Wilson,R.  
 Washu Xenopus EST project, 1999  
 Unpublished (1999)  
 Other-ESTs: d939c10.x1  
 Contact: Sandy Clifton, Ph.D.  
 Washu Xenopus EST project, 1999  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 Library constructed by Bruce Blumberg  
 Library normalized by Jihwan Song  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: Xenopus clone distribution information for  
 this library can be found through Research Genetics, visit their  
 web page at: <http://www.resgen.com/> Please reference the id listed  
 below when ordering this clone: Source lab clone id - xlnmg011e19  
 Seq primer: -40RP from Gibco  
 High quality sequence stop: 439.  
 Location/Qualifiers  
 1..442  
 /organism="Xenopus laevis"  
 /db\_xref="taxon:8155"  
 /clone\_id="IMAGE:3751290"  
 /clone\_id="Xenopus laevis gastrula non normalized"  
 /tissue\_type="gastrula (stages 10.5, 11.5 mixed)"  
 /lab\_host="Top-10 F"  
 /note="Vector: pBluescript SK-; Site: 1; EcoRI; Site 2:  
 XhoI; cDNA was prepared from 2ug of poly A+ RNA (equal  
 parts from stage 10.5 and stage 11.5 gastrulae).  
 EcoRI-XhoI cut cDNA was then ligated into unitap-XR  
 (Stratagene) with EcoRI at the 5' end and XhoI at the  
 3' end. SS-library phagemids were prepared by mass excision  
 from the original library and normalized by hybridization  
 to biotinylated driver (prepared from the same library by  
 PCR) to cot-omega of 11. After removal of hybrids and  
 excess driver by streptavidin sepharose chromatography,  
 the ss-phagemids were made double stranded and  
 electroporated into Top-10 F'. Original library  
 construction by Bruce Blumberg (Cho et al. 1991 Cell 67,  
 1111-1120). Note: This is a Xenopus Gene Collection (XGC)  
 library."

BASE COUNT 130 a 105 c 90 g 117 t  
 ORIGIN

Query Match 87.0%; Score 17.4; DB 11; Length 442;  
 Best Local Similarity 94.7%; Pred. No. 3.8e+02;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ccagttctcaagagcagc 19  
 |||  
 DB 54 CCATTCTCAAGAGCAGC 72

RESULT 4  
LOCUS B1102963 739 bp mRNA EST 26-JUN-2001  
DEFINITION 602888573p1 NCI\_CGAP\_Kid14 Mus musculus cDNA clone IMAGE:5043738  
5', mRNA sequence.  
ACCESSION B1102963  
VERSION B1102963.1 GI:14553856  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus

REFERENCE Eukaryota: Metazoa: Chordata: Cranialata: Vertebrata: Euteleostomi:  
Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Murinae: Mus.  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgaps-f@mail.nih.gov](mailto:cgaps-f@mail.nih.gov)  
Tissue Procurement: Jeffrey E. Green, M.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: L14M1120 row: a column: 19  
High quality sequence stop: 737.  
Location/Qualifiers

FEATURES  
source 1..739  
/organism="Mus musculus"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone\_lib="IMAGE:5043738"  
/clone\_lib="NCI\_CGAP\_Kid14"  
/lab\_host="DH10B (T1 phage-resistant)"  
/note="Organ: kidney; Vector: pCMV-Sport6; Site: 1; NCI:  
Site: 2; Salt: Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 1.75 kb. Constructed by Life  
Technologies. Note: this is a NCI\_CGAP Library. 1"  
Technologies. Note: this is a NCI\_CGAP Library. 1"

BASE COUNT 199 a 201 c 168 g 170 t 1 others  
ORIGIN  
Query Match 87.0%; Score 17.4; DB 11; Length 739;  
Best Local Similarity 94.7%; Pred. No. 3.9e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
2 cagttctcaaaagagcagc 20  
11111111111111111111  
202 CAGTCTCAAGCAGCATG 184

RESULT 5  
LOCUS CNS025H0 832 bp DNA GSS 12-MAY-2000  
DEFINITION Tetraodon nigroviridis genome survey sequence pUC-ori end of clone  
236j16 of library G from Tetraodon nigroviridis, genomic survey  
sequence.  
ACCESSION AL182061  
VERSION AL182061.1 GI:7820152  
KEYWORDS GSS; genome survey sequence.  
Tetraodon nigroviridis.  
Tetraodon nigroviridis.  
Tetraodon nigroviridis.  
Eukaryota: Metazoa: Chordata: Cranialata: Vertebrata: Euteleostomi:  
Actinopterygii: Neopterygii: Teleostei: Euteleostei: Neoteleostei:  
Acanthomorphi: Acanthopterygii: Percormorpha: Tetraodontiformes:  
Tetraodontidae: Tetraodon.  
1 (bases 1 to 832)  
Roest-Crollius, H., Jallion, O., Dasilva, C., Fizes, C., Fisher, C.,  
Bonneau, L., Billault, A., Quetier, F., Saurin, W., Bernot, A. and  
Weissenbach, J.  
TITLE Characterization and repeat analysis of the compact genome of the  
freshwater pufferfish Tetraodon nigroviridis

JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 832)  
AUTHORS Roest-Crollius, H., Jallion, O., Dasilva, C., Bonneau, L., Fisher, C.,  
Bernot, A., Fizes, C., Wincker, P., Brotier, P., Quetier, F.,  
Saurin, W. and Weissenbach, J.  
TITLE Human gene number estimate provided by genome wide analysis using  
Tetraodon nigroviridis DNA sequence  
JOURNAL Unpublished  
AUTHORS 3 (bases 1 to 832)  
REFERENCE Genoscope.  
TITLE Direct Submission  
JOURNAL Submitted (12-Apr-2000) to the EMBL/Genbank/DBJ databases  
COMMENT This sequence is a single read and was generated as part of a large  
scale clone-end sequencing project of the Tetraodon nigroviridis  
genome. For more information, please take a look at  
<http://www.genoscope.cns.fr/Tetraodon>.

FEATURES  
source 1..832  
/organism="Tetraodon nigroviridis"  
/db\_xref="taxon:99883"  
/clone\_lib="G"  
/clone\_lib="G"  
/note="Genoscope sequence ID : C0AG236DE08SP1-end :  
pUC-ori"  
pUC-ori

BASE COUNT 236 a 221 c 168 g 202 t 5 others  
ORIGIN

Query Match 87.0%; Score 17.4; DB 13; Length 832;  
Best Local Similarity 94.7%; Pred. No. 4e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
OY 1 ccagttctcaaaagagcagc 19  
11111111111111111111  
Db 362 CCAAGTCTCAAGCAGCATG 344

RESULT 6  
LOCUS AM176762 177 bp mRNA EST 16-NOV-1999  
DEFINITION RC0-CT0094-200899-002-B12 CT0094 Homo sapiens cDNA, mRNA sequence.  
ACCESSION AM176762  
VERSION AM176762.1 GI:6442799  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota: Metazoa: Chordata: Cranialata: Vertebrata: Euteleostomi:  
Mammalia: Eutheria: Primates: Catarrhini: Homiidae: Homo.  
1 (bases 1 to 177)  
HCCP <http://www.ludwig.org.br/ORESSES>.  
The FAPESP/LICR Human Cancer Genome Project  
Unpublished (1999)  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: [asimpson@ludwig.org.br](mailto:asimpson@ludwig.org.br)  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(<http://www.ludwig.org.br/scripts/gethtml2.pl?lib=RC0&t=RC0-CT0094-200899-002-B12&t3=1999-08-20&t4=1>)  
Seq primer: pUC 18 forward  
High quality sequence stop: 177.  
Location/Qualifiers

FEATURES  
source 1..177

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="CT0094"  
/dev\_stage="Adult"  
/note="Organ: colon; Vector: puc18; Site: 1; SmaI; Site: 2;  
freshwater pufferfish Tetraodon nigroviridis



ACCESSION AI927023 GI:5662987  
 VERSION AI927023.1  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 343)  
 NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP).  
 Tumor Gene Index  
 Unpublished (1997)  
 JOURNAL  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgapb@remail.nih.gov](mailto:cgapb@remail.nih.gov)  
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.  
 CDNA Library Preparation: M. Bento Soares, Ph.D.  
 DNA Sequencing by: Greg Lennon, Ph.D.  
 Clone distribution: NCI-CCAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: [www.bio.lnl.gov/btrp/image/image.html](http://www.bio.lnl.gov/btrp/image/image.html)  
 Insert length: 444 Std Error: 0.00  
 Seq primer: -40UP from Gibco.  
 Location/Qualifiers  
 1..343  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:2462290"  
 /clone\_lib="NCI-CCAP\_Kid11"  
 /lab\_host="DH10B"  
 /note="Organ: Kidney; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site\_1: Not I; Site\_2: Eco RI; Plasmid DNA from the normalized library NCI-CCAP\_Kid3 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clones 132376-132391, 145607-1456775, and 1500552-1502855). Subtraction by Bento Soares and M. Fatima Bonaldo."  
 BASE COUNT 87 a 78 c 97 g 80 t 1 others  
 ORIGIN  
 Query Match 84.0%; Score 16.8; DB 10; Length 343;  
 Best Local Similarity 90.0%; Pred. No. 7.1e+02;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 1 ccagttctcaagagagagag 20  
 ||||| ||||| ||||| |||||  
 Db 136 CCAGTCCCAAGACGACG 155  
 RESULT 10  
 LOCUS AW780736 350 bp mRNA EST 12-MAY-2000  
 DEFINITION s184f06.y1 Gm-cl037 Glycine max CDNA GENEOME SYSTEMS CLONE ID: Gm-cl037-732.5' Similar to TR-096305 Q96305 SIMILAR TO GLUTAREDOXIN ENCODED BY GENBANK ACCESSION NUMBER Z49699 ;, mRNA sequence.  
 ACCESSION AW780736  
 VERSION AW780736.1 GI:7795411  
 KEYWORDS EST.  
 SOURCE soybean.  
 ORGANISM Glycine max  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.  
 1 (bases 1 to 350)  
 Shoemaker,R., Keln,P., Vodkin,L., Erpelting,J., Corryell,V., Khanna,A., Holla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers

TITLE  
 JOURNAL  
 COMMENT Contact: Shoemaker R/Public Soybean EST Project  
 Public Soybean EST Project  
 Washington University School of Medicine  
 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: [est@wustl.edu](mailto:est@wustl.edu)  
 This clone is available through: Genome Systems, Inc. 4633 World Parkway Circle St. Louis, Missouri 63134 For further information call: (800) 430-0040 or (314) 427-3222 FAX: (888) 919-3324 or (314) 427-3424 or contact: [clones@genomesystems.com](mailto:clones@genomesystems.com) or [info@genomesystems.com](http://info@genomesystems.com) web site: [www.genomesystems.com](http://www.genomesystems.com)  
 putative full length read  
 vector to vector length is 351.  
 Location/Qualifiers  
 1..350  
 /organism="Glycine max"  
 /db\_xref="taxon:3847"  
 /clone="GENOME SYSTEMS CLONE ID: Gm-cl037-732"  
 /clone\_lib="Gm-cl037"  
 /issue\_type="Fully expanded leaves of greenhouse grown plants."  
 /dev\_stage="2 week old"  
 /lab\_host="DH10B"  
 /note="Vector: pSPORT1; Site\_1: NotI; Site\_2: SalI; This cDNA library was constructed from mRNA isolated from fully expanded leaves of greenhouse grown plants that were 2 weeks old. The library was prepared using the Life Technologies pSPORT1 cDNA library construction kit. Complementary DNA was synthesized from mRNA using a poly(dT) sequence with a NotI restriction site. SalI linkers/adapters were ligated to the blunt ended cDNA fragments followed by NotI digestion. The cDNA fragments were directionally cloned into the NotI-SalI restriction site of the pSPORT1 vector. The ligated cDNA fragments were transformed into E.coli Electro-Max DH10B host cells. This library was constructed in the laboratory of Dr. Jila Vodkin by Anu Khanna at the University of Illinois at Urbana-Champaign. email: [l.vodkin@uiuc.edu](mailto:l.vodkin@uiuc.edu)"  
 BASE COUNT 93 a 46 c 65 g 146 t  
 ORIGIN  
 Query Match 84.0%; Score 16.8; DB 10; Length 350;  
 Best Local Similarity 90.0%; Pred. No. 7.1e+02;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 1 ccagttctcaagagagagag 20  
 ||||| ||||| ||||| |||||  
 Db 102 CCAGTTCGCAAGACGACG 121  
 QY  
 RESULT 11  
 LOCUS N24917 356 bp mRNA EST 28-DEC-1995  
 DEFINITION y01h12.s1 Soares melanocyte 2NBHM Homo sapiens CDNA clone IMAGE:270023 3', mRNA sequence.  
 ACCESSION N24917  
 VERSION N24917.1 GI:1139067  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 356)  
 Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston



TITLE  
JOURNAL  
COMMENT

R. Williamson, A., Wohlmann, P. and Wilson, R.  
The Washu-Merck EST Project  
Unpublished (1995)  
Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@wustl.wustl.edu  
High quality sequence stops: 254  
Source: IMAGE Consortium, LNLN.  
This clone is available royalty-free through LNLN; contact the  
IMAGE Consortium (info@image.lnl.gov) for further information.  
Seq primer: ml3 -40 forward  
High quality sequence stop: 254.

## FEATURES

source

1. 356  
/organism="Homo sapiens"  
/db\_xref="GDB:3879665"  
/db\_xref="taxon:9606"  
/clone="IMAGE:270023"  
/clone\_lib="Soares melanocyte 2NBHM"  
/sex="Male"  
/tissue="type="melanocyte"  
/lab\_host="DH10B (ampicillin resistant)"  
/note="Vector: pT73D (pharmacia) with a modified  
polylinker. Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA  
was primed with a Not I - c1190(dT) primer (5'  
TGTTACATCTGAGTGGAGCGCGCGACTTTTTTTTTTTT 3'),  
double-stranded cDNA was size selected, ligated to Eco RI  
adapters (pharmacia), digested with Not I and cloned into  
the Not I and Eco RI sites of a modified pT73 vector  
(pharmacia). Library constructed by Bento Soares and  
M. Fatima Bonaldo. RNA from normal foreskin melanocytes  
(FS34) was kindly provided by Dr. Anthony P. Albino."

## BASE COUNT

90 a 81 c 103 g 82 t

Query Match 84.0%; Score 16.8; DB 11; Length 356;  
Best Local Similarity 90.0%; Pred. No. 7.1e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 ccaagttcgaagagcagc 20  
||||| |||||||||  
DB 138 CCAAGTCCCAAGAGCAGC 157

## R T 12

LOCUS AM015308 358 bp mRNA EST 10-SEP-1999  
DEFINITION UT-H-B10-aap-e-04-0-UL.s1 NCLCGAP\_Sub1 Homo sapiens cDNA clone  
IMAGE:2709991.3, mRNA sequence.

ACCESSION AM015308  
VERSION AM015308.1 GI:5864065  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index

JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cga@bbs-remail.nih.gov

The sequence contained an oligo-dT track that was present in the  
oligonucleotide that was used to prime the synthesis of first  
strand cDNA and therefore this may represent a bonafide poly A  
tail. cDNA Library Preparation: M.B. Soares Lab Clone distribution:  
NCI-CGAP clone distribution information can be found through the  
I.M.A.G.E. Consortium/LNLN at:

## FEATURES

source

www.bio.lnl.gov/Db/rlp/image/image.html  
Seq primer: M13 forward  
PolyA=yes.

Location/Qualifiers  
1. 358

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2709991"  
/clone\_lib="NCI-CGAP\_Sub1"  
/lab\_host="DH10B (Life Technologies)"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified  
polylinker. Site\_1: Not I; Site\_2: Eco RI; The  
NCI-CGAP\_Sub1 library is a subtracted library derived from  
BL. It constitutes a mixture of 21 normalized or  
subtracted NCI-CGAP libraries: NCI-CGAP\_C04,  
NCI-CGAP\_Pr22, NCI-CGAP\_Pr28, NCI-CGAP\_C010, NCI-CGAP\_C016  
NCI-CGAP\_Kid1, NCI-CGAP\_Lym2, NCI-CGAP\_Br23, NCI-CGAP\_Lu5  
NCI-CGAP\_Kid11, NCI-CGAP\_Lu24, NCI-CGAP\_Lu19, NCI-CGAP\_GC4, NCI-CGAP\_GC6  
, NCI-CGAP\_Brn25. These 21 libraries were pooled and a  
single-stranded DNA preparation of the resulting mixture  
was used as a tracer in a subtractive hybridization with a  
driver whose composition is detailed below: NCI-CGAP\_Kid3  
pool 1 LLM 3334-3337, 3682-3683, 3798-3803 (IMAGE  
Clonoids 1322376-1323911, 1456008-1456775, 1500552-1502855  
) NCI-CGAP\_Kid5 pool 1 LLM 3338-3342, 3732-3725.  
3776-3778 (IMAGE Clonoids 1323912-1325831,  
1471368-1472903, 1492104-1493255) NCI-CGAP\_Lu5 pool 1 LLM  
3575-3582, 3851-3854 (IMAGE Clonoids 1414920-1417991,  
1520904-1522439) NCI-CGAP\_GC4 pool 1 LLM 3164-3167,  
3716-3720, 3733-3735 (IMAGE Clonoids 1257096-1258631,  
1469064-1470983, 1475592-1476743) NCI-CGAP\_Pr22 pool 1  
LLM 2457-2459, 2758-2759, 3062-3068 (IMAGE Clonoids  
985608-986759, 1101192-1101928, 1217928-1220615)  
NCI-CGAP\_C010 pool 1 LLM 2644-2653, 2871-2872 (IMAGE  
Clonoids 1057416-1061255, 1144584-1145351) The resulting  
subtracted library contained 530,000 recombinants.  
Subtraction was performed as previously described [Bonaldo  
, Lennon & Soares (1996)]: Normalization and Subtraction:  
Two Approaches To Facilitate Gene Discovery. Genome  
Research 6, 791-806.  
TAG LIB-NCI CGAP Pr22  
TAG TISSUE-Prostate  
TAG\_SEQ-AAAGTC\*

BASE COUNT 85 a 79 c 96 g 98 t

Query Match 84.0%; Score 16.8; DB 10; Length 358;  
Best Local Similarity 90.0%; Pred. No. 7.1e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 ccaagttcgaagagcagc 20  
||||| |||||||||  
DB 156 CCAAGTCCCAAGAGCAGC 175

## RESULT 13

LOCUS C84007 359 bp mRNA EST 28-APR-1999  
DEFINITION C84007 Dictyostellium discoideum SS (H. Urushihara) Dictyostellium  
discoideum cDNA clone SSC227, mRNA sequence.

ACCESSION C84007  
VERSION C84007.1 GI:2723613  
KEYWORDS EST.  
SOURCE Dictyostellium discoideum.  
ORGANISM Dictyostellium discoideum

REFERENCE Eukaryota; Mycetozoa; Dictyosteliales; Dictyostellium.  
AUTHORS 1 (bases 1 to 359)  
Morio, T., Urushihara, H., Saito, T., Ugawa, Y., Mizuno, H., Yoshida, M.,  
Yoshino, R., Mita, B.N., Pi, M., Sato, T., Takemoto, K., Yasukawa, H.,  
Williams, J., Maeda, M., Takeuchi, I., Ochiai, H. and Tanaka, Y.

```

"organism":"Homo sapiens"
"/db_xref":"taxon:9606"
"clone IMAGE:2062902"
"clone_lib":"NCI CGAP C616"
"tissue_type":"colon tumor, RER+"
"host":"DH10B"
"note":"organ: colon; Vector: pF73D-pac (pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI;

```

BASE COUNT	113 a	84 c	80 g	102 t
ORIGIN				

Query Match 84.0%; Score 16.8; DB 13; Length 380;  
 Best Local Similarity 90.0%; Pred. No. 7.2e+02;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 ccagttctcaagagcagc 20.  
 |||||  
 Db 123 CCACTTCTCAAGTACCAGG 142

Search completed: March 4, 2002, 02:14:35  
 Job time: 6484 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 4, 2002, 05:21:18 : Search time 578.64 Seconds  
(without alignments)  
29,632 Million cell updates/sec

Title: US-08-956-991-5

Perfect score: 20  
Sequence: 1 ccagttcacaagagcagc 20

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq\_1101:\*

1: /SID52/gcgdata/geneseq/geneseqn/NA1980.DAT.\*  
2: /SID52/gcgdata/geneseq/geneseqn/NA1981.DAT.\*  
3: /SID52/gcgdata/geneseq/geneseqn/NA1982.DAT.\*  
4: /SID52/gcgdata/geneseq/geneseqn/NA1983.DAT.\*  
5: /SID52/gcgdata/geneseq/geneseqn/NA1984.DAT.\*  
6: /SID52/gcgdata/geneseq/geneseqn/NA1985.DAT.\*  
7: /SID52/gcgdata/geneseq/geneseqn/NA1986.DAT.\*  
8: /SID52/gcgdata/geneseq/geneseqn/NA1987.DAT.\*  
9: /SID52/gcgdata/geneseq/geneseqn/NA1988.DAT.\*  
10: /SID52/gcgdata/geneseq/geneseqn/NA1989.DAT.\*  
11: /SID52/gcgdata/geneseq/geneseqn/NA1990.DAT.\*  
12: /SID52/gcgdata/geneseq/geneseqn/NA1991.DAT.\*  
13: /SID52/gcgdata/geneseq/geneseqn/NA1992.DAT.\*  
14: /SID52/gcgdata/geneseq/geneseqn/NA1993.DAT.\*  
15: /SID52/gcgdata/geneseq/geneseqn/NA1994.DAT.\*  
16: /SID52/gcgdata/geneseq/geneseqn/NA1995.DAT.\*  
17: /SID52/gcgdata/geneseq/geneseqn/NA1996.DAT.\*  
18: /SID52/gcgdata/geneseq/geneseqn/NA1997.DAT.\*  
19: /SID52/gcgdata/geneseq/geneseqn/NA1998.DAT.\*  
20: /SID52/gcgdata/geneseq/geneseqn/NA1999.DAT.\*  
21: /SID52/gcgdata/geneseq/geneseqn/NA2000.DAT.\*  
22: /SID52/gcgdata/geneseq/geneseqn/NA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	20	100.0	20	19	AAV31983 Human Down syndrome
2	20	100.0	6413	19	AAV31988 Human Down syndrome
3	20	100.0	6604	19	AAV31981 Human Down syndrome
4	18.4	92.0	2173	19	AAV31987 Mouse Down syndrome
5	18.4	92.0	54548	21	AAZ45596 DNA sequence of th
6	17.4	87.0	2197	21	AAZ89617 A. sydowi afl-SST
7	16.8	84.0	187	16	AAZ22739 Human gene signal
8	16.4	82.0	12923	10	AAAG0338 Sequence of human
9	15.8	79.0	209	22	AAI25968 Probe #15901 for g
10	15.8	79.0	209	22	AAI53359 probe #22045 used
11	15.8	79.0	563	22	AAI16773 Probe #6706 for ge

C	12	15.8	79.0	563	22	AAI40280	Probe #8966 used t
	13	15.8	79.0	564	22	AAI48672	Probe #17358 used t
	14	15.8	79.0	564	22	AAI08987	Probe #8978 used t
	15	15.8	79.0	673	22	AAH00627	Enterococcus avium
	16	15.8	79.0	1972	22	AAI35523	Probe #4209 used t
	17	15.8	79.0	1972	22	AAI03989	Probe #3980 used t
	18	15.8	79.0	2947	22	AAH18697	Human CDNA sequenc
	19	15.8	79.0	5406	19	AAV52299	Streptococcus pneu
	20	15.8	79.0	530	22	AAH13400	Human CDNA clone (
	21	15.4	77.0	796	22	AAH05981	Human CDNA clone (
	22	15.4	77.0	1665	22	AAH14284	Human CDNA sequenc
	23	15.4	77.0	4646	20	AAZ20577	Polynucleotide seq
	24	15.4	77.0	4949	21	AAZ69899	Human ORFX ORF2544
	25	15.2	76.0	47	21	AAZ68502	Human map-related
	26	15.2	76.0	167	22	AAH09066	Human CDNA clone (
	27	15.2	76.0	180	22	AAI51608	Probe #20294 used
	28	15.2	76.0	351	21	AAZ4771	Human ORFX ORF326
	29	15.2	76.0	478	22	AAI38475	Probe #7161 used t
	30	15.2	76.0	750	22	AAI59635	Human polynucleotl
	31	15.2	76.0	798	22	AAI57849	Human polynucleotl
	32	15.2	76.0	892	21	AAZ49173	CDNA encoding huma
	33	15.2	76.0	1209	20	AAZ81466	Mosaic fla gene c
	34	15.2	76.0	1228	20	AAZ81465	Mosaic fla gene c
	35	15.2	76.0	1229	20	AAZ81464	Nucleotide sequenc
	36	15.2	76.0	1402	21	AAZ6896	Zea mays DNA fragm
	37	15.2	76.0	1407	22	AAZ42469	Human cDNA encodin
	38	15.2	76.0	1408	21	AAZ43700	Human transport pr
	39	15.2	76.0	1881	22	AAZ2645	Human cDNA encodin
	40	15.2	76.0	1881	22	AAZ27735	Human cDNA sequenc
	41	15.2	76.0	1930	22	AAZ22881	Human immunoglobul
	42	15.2	76.0	1996	22	AAH15280	Human cDNA sequenc
	43	15.2	76.0	2303	22	AAZ30955	Human cDNA sequenc
	44	15.2	76.0	2336	22	AAH16558	Human secreted pro
	45	15.2	76.0	2652	21	AAZ61272	

#### ALIGNMENTS

RESULT 1

AAV31983

ID AAV31983 standard, cDNA: 20 BP.

AC AAV31983:

XX 28-SEP-1998 (first entry)

XX DT

XX DE Human Down syndrome-cell adhesion molecule US-CAM primer B9-131F.

XX AC

XX DS-CAM1: Down syndrome-cell adhesion molecule; neural cell;

KW signal transduction; trisomy 21; mental retardation;

KW holoprosencephaly; corpus callosum agenesis;

KW schizencephaly; diagnosis; assay; human; PCR; primer; ss.

XX OS Synthetic.

XX OS Homo sapiens.

XX PN WC0817795-A1.

XX PD 30-APR-1998.

XX PF 23-OCT-1997; 97WD-US19547.

XX PR 25-OCT-1996; 96US-0029322.

XX PA (CEDA-) CEUAKS SINAI MEDICAL CENT.

XX PI Korenberg JK;

XX PT WPI: 1998-271791/24.

XX DR New isolated Down's Syndrome-cell adhesion molecule - used to

XX PT develop products for detection, diagnosis and therapy of

developmental and neurological abnormalities

Example 5; Page 79; 109pp; English.

Forward primer B9-131F and reverse primer B9-131R (see AAV31984)

were used in RT-PCR assays of cDNA libraries of various human

tissues to determine human Down syndrome-cell adhesion molecule

DS-CAM (see AAV31981 and AAV31988) expression. The results

demonstrated expression of human DS-CAM mRNA in foetal and adult

brain, and foetal kidney. In addition, a breast carcinoma cell

line showed expression of human DS-CAM mRNA. DS-CAM polypeptides

(see AAW42086-87) are associated with developmental and neurological

processes. The polypeptides and nucleic acids are used to develop

products for the detection, diagnosis and therapy of developmental

and neurological abnormalities.

Sequence 20 BP; 6 A; 5 C; 6 G; 3 T; 0 other;

Very Match 100.0%; Score 20; DH 19; Length 20;  
Local Similarity 100.0%; Pred. No. 1.6;  
Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ccagctcrraaaggacagcagg 20  
|||||  
1 ccagtlctcaaggagcagcagg 20

RESULT 2  
AAV31988  
ID AAV31988 standard; cDNA; 6413 BP.  
XX  
AC AAV31988;  
XX  
DT 28-SEP-1998 (first entry)  
XX  
DE Human Down syndrome-cell adhesion molecule DS-CAM2 cDNA.  
XX  
DS DS-CAM2; Down syndrome-cell adhesion molecule; neural cell;  
XX signal transduction; trisomy 21; mental retardation;  
KM holoprosencephaly; corpus callosum agenesis;  
KW schizencephaly; diagnosis; assay; human; ds; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 453..5168  
FT /\*tag= a  
X/MO9817795-A1.  
XX  
PD 30-APR-1998.  
XX  
PF 23-OCT-1997; 97MO-US19547.  
XX  
PK 25-OCT-1996; 96US-0029322.  
XX  
PA (CEDA-) CERARS SINAI MEDICAL CENT.  
XX  
PI Korenberg JR.  
XX  
DR WPL: 1998-271791/24.  
DX P-PSDB: AAW42087.  
XX  
PT New isolated Down's Syndrome-cell adhesion molecule - used to  
PT develop products for detection, diagnosis and therapy of  
PT developmental and neurological abnormalities  
XX  
PS Claim 3; Page 83-90; 109pp; English.  
XX  
CC This cDNA clone codes for Down syndrome-cell adhesion molecule  
CC DS-CAM2 (see AAW42087), an extracellular, soluble protein belonging  
CC to a novel subclass of the Ig superfamily with highest homology to

Query Match	Best Local	Similarity	Score	20:	DB	19:	Length	6413:
Matches	20:	Conservative	100.0%	0:	Mismatches	0:	Indels	0:
Gaps	0:							
Oy	1	ccagttctcaaaagagcagc	20					
Db	4877	ccagttctcaaaagagcagc	4896					
RESULT	3							
ID	AAV31981	standard: cDN/	6604	BP.				
XX	XX							
AC	AAV31981:							
DT	28-SEP-1998	(first entry)						
XX	XX							
DE	Human Down syndrome-cell adhesion molecule DS-CAM1 cDNA.							
XX	XX							
KM	DS-CAM1: Down syndrome-cell adhesion molecule: neural cell:							
KM	signal transduction; trisomy 21; mental retardation;							
KW	holoprosencephaly; corpus callosum; spina; spina;							
XX	schizencephaly; diagnosis; assay; human; ds; ss.							
OS	Homo sapiens.							
XX	XX							
PH	Key	Location/Qualifiers						
FT	CDS	453..6185						
FT		/*lag- a						
FT	sig_peptide	452..521						
FT		/*lag- b						
FT	mat_peptide	522..6162						
FT		/*lag- c						
XX	XX							
PN	W09817/95-A1.							
XX	XX							
PD	30-APR-1998.							
XX	XX							
PF	23-OCT-1997:	97WO-US19547.						
XX	XX							
PR	25-OCT-1996:	96US-0029322.						
XX	XX							
PA	(CEDA-) CEDAKS SINAI MEDICAL CENT.							
XX	XX							
PI	Korenberg JK:							
XX	XX							
DR	WPI: 1998-271791/24.							

UK P-PSDB: AAM42086.  
 XX New isolated Down's Syndrome-cell adhesion molecule - used to  
 PT develop products for detection, diagnosis and therapy of  
 PT developmental and neurological abnormalities  
 XX  
 PS Claim 3: Page 65-72; 109pp; English.

CC This cDNA clone codes for Down syndrome-cell adhesion molecule  
 CC DS-CAM1 (see AAM42086), a cell surface glycoprotein belonging to a  
 CC novel subclass of the Ig superfamily with highest homology to  
 CC neural cell adhesion molecules. A modified direct cDNA selection  
 CC technique was applied to bacterial and plasmid artificial chromosomes  
 CC between ET52 and MX1 by using cDNA from trisomy 21 human foetal  
 CC brain. A unique cDNA fragment, designated B31 (see AAV31982), was  
 CC obtained and used to screen a trisomy 21 human foetal brain (14-wk  
 CC gestation) cDNA library. Further clones were obtained by exon  
 CC trapping. The complete DS-CAM1 cDNA sequence was deduced from  
 CC overlapping clones. A splice variant cDNA (see AAV31988), encoding  
 CC non-membrane bound DS-CAM2 (see AAM42087), was also identified. The  
 CC DS-CAM gene spans 900-1200 kb of genomic DNA and maps at chromosome  
 CC 21q22.2-22.3. The invention also provides murine DS-CAM partial  
 CC sequences (see AAV31985-87), expression vectors and host cells,  
 CC transgenic animals, antisense oligonucleotides, and primers useful  
 CC for amplification of DS-CAM nucleic acid. DS-CAM polypeptides are  
 CC associated with developmental and neurologic processes. They can  
 CC be used in e.g. neural prosthetic devices used in enervation  
 CC methods of repairing (regenerating) damaged or severed peripheral  
 CC nerves. The products can also be used in detection, diagnosis and  
 CC therapy of developmental and neurological abnormalities such as  
 CC Down syndrome, mental retardation, holoprosencephaly, agenesis of  
 CC the corpus callosum, or schizencephaly. Antisense oligonucleotides  
 CC are used for inhibiting translation of mRNA.

Sequence 6604 BP; 1673 A; 1827 C; 1768 G; 1336 T; 0 other;

Query Match 100.0%; Score 20; DB 19; Length 6604;  
 Best Local Similarity 100.0%; Pred. No. 2.7;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

YY 1 ccaagttctcaaaagagcagc 20  
 |||||  
 Db 4877 ccaagttctcaaaagagcagc 4896

RESULT 4

AAV31987 standard; cDNA: 2173 BP.

AAV31987:

28-SEP-1998 (first entry)

Mouse Down syndrome-cell adhesion molecule DS-CAM 3' cDNA.

DS-CAM: Down syndrome-cell adhesion molecule; neural cell;

Signal transduction; trisomy 21; mental retardation;

holoprosencephaly; corpus callosum agenesis;

schizencephaly; diagnosis; assay; mouse; ds; ss.

Mus sp.

MO9817795-A1.

30-APR-1998.

23-OCT-1997; 97MO-US19547.

25-OCT-1996; 96US-0029322.

(CEDA-) CEDARS SINAI MEDICAL CENT.

PI Korenberg JK;  
 XX WPI: 1998-271791/24.

XX New isolated Down's Syndrome-cell adhesion molecule - used to  
 PT develop products for detection, diagnosis and therapy of  
 PT developmental and neurological abnormalities  
 XX  
 PS Claim 2: Page 81-83; 109pp; English.

CC This cDNA sequence comprises the 3' region of a cDNA clone  
 CC for murine Down syndrome-cell adhesion molecule (DS-CAM), a  
 CC member of a novel subclass of the Ig superfamily with homology to  
 CC neural cell adhesion molecules. The middle region (see AAV31986)  
 CC and 5' region (see AAV31987) of the clone are also provided. The  
 CC murine DS-CAM clone was isolated from a C57 Black/6 mouse brain  
 CC cDNA library using human DS-CAM cDNA (see AAV31981 and AAV31982) as  
 CC probe. The invention also provides human DS-CAM proteins (see  
 CC AAM42086-87), as well as expression vectors and host cells,  
 CC transgenic animals, antisense oligonucleotides, and primers useful  
 CC for amplification of DS-CAM nucleic acids. DS-CAM polypeptides are  
 CC associated with developmental and neurological processes. They can  
 CC be used in e.g. neural prosthetic devices used in enervation  
 CC methods of repairing (regenerating) damaged or severed peripheral  
 CC nerves. The products can also be used in detection, diagnosis and  
 CC therapy of developmental and neurological abnormalities such as  
 CC Down syndrome, mental retardation, holoprosencephaly, agenesis of  
 CC the corpus callosum, or schizencephaly. Antisense oligonucleotides  
 CC are used for inhibiting translation of mRNA.

Sequence 2173 BP; 585 A; 598 C; 554 G; 436 T; 0 other;

Query Match 92.0%; Score 18.4; DB 19; Length 2173;  
 Best Local Similarity 95.0%; Pred. No. 14;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

YY 1 ccaagttctcaaaagagcagc 20  
 |||||  
 Db 479 ccaagttctcaaaagagcagc 498

RESULT 5

AA245596 standard; DNA: 54548 BP.

AA245596:

06-APR-2000 (first entry)

DNA sequence of the wild type human PCY64 region of Xq13.

Human; PCY64 region; X chromosome; q13 region; polymorphism;

mental retardation; autism; depression; bipolar affective disorder;

hypothyroidism; OPA gene; neuropsychiatric disorder; ss.

Homo sapiens.

Location/Qualifiers

key exon

promoter

promoter

exon

note="human OPA promoter"

note="human OPA promoter"

note="human OPA promoter"

```
FT exon 2321..2512 /*tag= e
FT /number= 3
FT /note= "human OPA promoter"
FT exon 2649..2805 /*tag= f
FT /number= 4
FT /note= "human OPA promoter"
FT exon 3608..3787 /*tag= g
FT /number= 5
FT /note= "human OPA promoter"
FT exon 3962..4072 /*tag= h
FT /number= 6
FT /note= "human OPA promoter"
FT exon 4197..4451 /*tag= i
FT /number= 7
FT /note= "human OPA promoter"
FT exon 4839..4986 /*tag= j
FT /number= 8
FT /note= "human OPA promoter"
FT exon 5148..5247 /*tag= k
FT /number= 9
FT /note= "human OPA promoter"
FT exon 5377..5513 /*tag= l
FT /number= 10
FT /note= "human OPA promoter"
FT exon 5734..5865 /*tag= m
FT /number= 11
FT /note= "human OPA promoter"
FT exon 6233..6359 /*tag= n
FT /number= 12
FT /note= "human OPA promoter"
FT exon 6798..7027 /*tag= o
FT /number= 13
FT /note= "human OPA promoter"
FT exon 7403..7483 /*tag= p
FT /number= 14
FT /note= "human OPA promoter"
FT exon 7615..7784 /*tag= q
FT /number= 15
FT /note= "human OPA promoter"
FT exon 7989..8133 /*tag= r
FT /number= 16
FT /note= "human OPA promoter"
FT exon 8301..8351 /*tag= s
FT /number= 17
FT /note= "human OPA promoter"
FT exon 8674..8792 /*tag= t
FT /number= 18
FT /note= "human OPA promoter"
FT exon 8979..9122 /*tag= u
FT /number= 19
FT /note= "human OPA promoter"
FT exon 9606..9769 /*tag= v
FT /number= 20
FT /note= "human OPA promoter"
FT exon 9973..10104 /*tag= w
FT /number= 21
FT /note= "human OPA promoter"
FT exon 10300..10757 /*tag= x
FT /number= 22
FT /note= "human OPA promoter"
FT exon 10933..11077 /*tag= y
FT /number= 23
FT /note= "human OPA promoter"
FT exon 11235..11355 /*tag= z
FT /number= 24
FT /note= "human OPA promoter"
FT exon 11751..12065 /*tag= aa
FT /number= 25
FT /note= "human OPA promoter"
FT exon 12317..12492 /*tag= ab
FT /number= 26
FT /note= "human OPA promoter"
FT exon 12672..12851 /*tag= ac
FT /number= 27
FT /note= "human OPA promoter"
FT exon 14187..14258 /*tag= ad
FT /number= 28
FT /note= "human OPA promoter"
FT exon 14710..14843 /*tag= ae
FT /number= 29
FT /note= "human OPA promoter"
FT exon 15014..15175 /*tag= af
FT /number= 30
FT /note= "human OPA promoter"
FT exon 15487..15598 /*tag= ag
FT /number= 31
FT /note= "human OPA promoter"
FT exon 15765..15854 /*tag= ah
FT /number= 32
FT /note= "human OPA promoter"
FT exon 16999..17109 /*tag= ai
FT /number= 33
FT /note= "human OPA promoter"
FT exon 17351..17492 /*tag= aj
FT /number= 34
FT /note= "human OPA promoter"
FT exon 17736..17897 /*tag= ak
FT /number= 35
FT /note= "human OPA promoter"
FT exon 18927..19301 /*tag= al
FT /number= 36
FT /note= "human OPA promoter"
FT exon 19525..19675 /*tag= am
FT /number= 37
FT /note= "human OPA promoter"
FT exon 19833..20029 /*tag= an
FT /number= 38
FT /note= "human OPA promoter"
FT exon 20207..20281 /*tag= ao
```

```

1
FT /number= 39
FT /note= "human OPA promoter"
FT 20372..20590
FT /*tag= ap
FT /number= 40
FT /note= "human OPA promoter"
FT 23283..23505
FT /*tag= aq
FT /number= 41
FT /note= "human OPA promoter"
FT 23878..24018
FT /*tag= at
FT /number= 42
FT /note= "human OPA promoter"
FT 24531..24612
FT /*tag= as
FT /number= 43
FT /note= "human OPA promoter"
FT 24823..25096
FT /*tag= at
FT /number= 44
FT /note= "human OPA promoter"
FT 26123..26275
FT /*tag= au
FT /note= "probable NL-3 promoter"

Query Match 92.0%; Score 18.4; DB 21; Length 54548;
Best Local Similarity 95.0%; Pred. No. 19;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 1 ccaagttcctcaagaagacagc 20
    ||||| ||||| ||||| |||||
Db 28196 ccaagttcctcaagaagacagc 28215

```

```

RESULT 6
AAZ89617/c
ID AAZ89617 standard; cDNA: 2197 BP.
XX
AC AAZ89617;
XX
DT 23-JUN-2000 (first entry)
XX
DE A. sydowi afl-SST cDNA.
XX
KM Fructosyl transferase; afl-SST; polyfructose; surfactant manufacture;
viscosity; sedimentation; food; ss.
XX
KW Aspergillus sydowi.

```

```

A.. Location/Qualifiers
FH Key 21..2069
FT CDS /tag= a
FT /product= "afl-SST"
XX
XX DE19840028-A1.
XX
XX PD 09-MAR-2000.
XX
XX PF 02-SEP-1998; 98DE-1040028.
XX
XX PR 02-SEP-1998; 98DE-1040028.
XX
XX PA (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
XX
XX PI Heyer A, Reim J, Wendenburg R;
XX
XX DK WPI: 2000-257741/23.
XX
XX DR P-PSDB: AAT51947.
XX
XX Novel DNA encoding Aspergillus fructosyl transferase useful for
PT production of polyfructose which is used to manufacture surfactants and
PT suspension agents

```

```

XX
PS Claim 1b: Page 12-15; 28pp; German.
XX
XX This invention describes a novel nucleic acid molecule (I) encoding a
CC fructosyl transferase isolated from Aspergillus sydowi. The fructosyl
CC transferase is useful for the production of polyfructose. Polyfructose
CC can be used to manufacture surfactants, to increase the viscosity of an
CC aqueous system, as a suspending agent, to accelerate sedimentation and
CC in the formation of complexes or to bind water. The host cell can be
CC used as a food. This sequence encodes the A. sydowi fructosyl transferase
CC designated afl-SST which is described in the method of the invention.
XX
SO Sequence 2197 BP; 515 A; 587 C; 575 G; 520 T; 0 other;

```

```

Query Match 87.0%; Score 17.4; DB 21; Length 2197;
Best Local Similarity 94.7%; Pred. No. 42;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 1 ccaagttcctcaagaagacagc 19
    ||||| ||||| ||||| |||||
Db 728 CCAGTTCCTCAAGAGAGAG 710

```

```

RESULT 7
AAT22739/c
ID AAT22739 standard; cDNA to mRNA: 187 BP.
XX
XX AAT22739;
XX
XX 14-AUG-1996 (first entry)
XX
DE Human gen. signature H0KGS04398.
XX
XX Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KM human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
XX
XX Homo sapiens.
XX
XX WO9514772-A1.
XX
XX PD 01-JUN-1995.
XX
XX PF 11-NOV-1994; 94KW-100194b.
XX
XX PR 12-NOV-1994; 94JP-045504.
XX
XX PA (MATS/) MATSUHARA K.
XX
XX PI (OKUH/) OKUHO K.
XX
XX Matsubara K, Okubo K;
XX
XX WPI: 1995-206931/27.
XX
XX Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
XX
XX Claim 1: Page 1203; 2245pp; Japanese.
XX
XX A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in AAT19001-126837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared from
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.

```



CC The appearance frequency of a given GS in a cDNA library can be  
 CC determined (esp. using primers and probes derived from the GS  
 CC sequences) as a means of diagnosing abnormal cell function or for  
 CC recognising different cell types.  
 XX  
 SQ Sequence 187 BP: 48 A; 42 C; 42 G; 49 T; 6 other;

Query Match 84.0%; Score 16.8; DB 16; Length 187;  
 Best Local Similarity 90.0%; Pred. No. 64;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 ccagttctcaagagcagcagg 20  
 ||||| ||||| ||||| |||||  
 Db 46 CCAGTCCCAAGAGCAGG 27

RESULT 8  
 AAN90338/C  
 17 AAN90338 standard; cDNA: 12923 BP.

AAN90338;

29-MAR-1992 (first entry)

Sequence of human muscular dystrophy (MD) cDNA.

Dystrophin: muscular dystrophy; probe; antibody; diagnosis;  
 prenatal; heterozygote; gene therapy; genetic screening;  
 foetal screening; ss.

Homo sapiens.

Key Location/Qualifiers  
 CDS 209..12923  
 /\*tag= a

W08906286-A.

13-JUL-1989.

16-DEC-1988; 88WO-US04504.

22-DEC-1987; 87US-0136618.

(CHIL-) CHILDRENS MED CENT.

Kunkel LM, Monaco A, Hoffman EP, Koenig M;

WPI; 1989-220587/30.

P-PSDB; AAP90373.

Muscular dystrophy gene - used for prepn. of probes, dystrophin  
 polypeptide and antibodies for diagnosis and therapy of muscular  
 dystrophy

Disclosure: Fig 5; 68pp: English.

The inventors claim an MD probe comprising a purified ss NA-SQ which  
 hybridises to at least a part of the MD gene; pure dystrophin (DS)  
 polypeptide, purified NA encoding DS and antibodies (Ab) to DS. The  
 CC probes are equal to or greater than 10b of one of 12 cDNA sequences  
 CC deposited as ATCC 58666-57677. The MD gene is human, or a murine Dmd  
 CC gene.

Sequence 12923 BP: 4296 A; 2613 C; 2963 G; 3044 T; 7 other;

Query Match 82.0%; Score 16.4; DB 10; Length 12923;  
 Best Local Similarity 94.4%; Pred. No. 1.5e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 2 ccagttctcaagagcagcagg 19

Db 11904 CACTTCTCAATGACGAG 11887  
 ||||| ||||| ||||| |||||

RESULT 9  
 AAI25968/C  
 ID AAI25968 standard; DNA: 209 BP.

AAI25968;

12-OCT-2001 (first entry)

Probe #15901 for gene expression analysis in human cervical cell sample.

Probe: human; microarray; gene expression; cervical epithelial cell;  
 cervical cancer; ss.

Homo sapiens.

W0200157278-A2.

09-AUG-2001.

30-JAN-2001; 2001WO-US00670.

04-FEB-2000; 2000US-0180312.

26-MAY-2000; 2000US-0207456.

30-JUN-2000; 2000US-0608408.

03-AUG-2000; 2000US-0632366.

21-SEP-2000; 2000US-0234687.

27-SEP-2000; 2000US-0236359.

04-OCT-2000; 2000GB-0024263.

(MOLE-) MOLECULAR DYNAMICS INC.

Penn SG, Hanzel DK, Chen W, Rank DR.

WPI; 2001-488901/53.

Human genome-derived single exon nucleic acid probes useful for  
 analyzing gene expression in human cervical epithelial cells -

Claim 25; SEQ ID NO 15901; 487pp; English.

The present invention relates to human single exon nucleic acid probes

(SENP). The present sequence is one such probe. The SENPs are derived

from human HeLa cells. The SENPs can be used to produce a single exon

sample derived from human cervical epithelial cells. By measuring gene

expression, the probes are therefore useful in grading and/or staging

CC of diseases of the cervix, notably cervical cancer.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

at ftp.wipo.int/pub/published\_pct\_sequences.

Sequence 209 BP: 47 A; 59 C; 54 G; 49 T; 0 other;

Oy 1 ccagttctcaagagcagcagg 19  
 ||| ||||| ||||| |||||

Db 175 CCGGCTCTCAAGAGCAGCAG 157

Query Match 79.0%; Score 15.8; DB 22; Length 209;  
 Best Local Similarity 89.5%; Pred. No. 1.9e+02;  
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

RESULT 10  
 AAI53359/C  
 ID AAI53359 standard; DNA: 209 BP.

AAI53359;



PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI; 2001-488897/53.  
XX Human genome-derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human placenta  
XX  
XX Claim 25: SEQ ID NO 8966; 654bp; English.  
XX  
CC The present invention relates to single exon nucleic acid probes (SENPs).  
CC The present sequence is one such probe. The probes are useful for  
CC producing a microarray for predicting, measuring and displaying gene  
CC expression in samples derived from human placenta. The probes are useful  
CC for antenatal diagnosis of human genetic disorders.  
XX  
SQ Sequence 563 BP; 115 A; 153 C; 178 G; 117 T; 0 other;

Query Match 79.0%; Score 15.8; DB 22; Length 563;  
Best Local Similarity 89.5%; Pred. No. 2.1e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 ccagttctcaagaagcag 19  
11111111111111111111  
DB 326 CCGGCTCTCAAGAGAGAG 308

## RESULT 13

AA148672  
ID AA148672 standard; DNA: 564 BP.

AC AA148672;

DT 17-OCT-2001 (first entry)

XX Probe #17358 used to measure gene expression in human placenta sample.

DE Probe: microarray; human; placenta; antenatal diagnosis;

KW Probe: microarray; human; placenta; antenatal diagnosis;

KW genetic disorder; ss.

OS Homo sapiens.

PN WO200157272-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00663.

PS 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX

XX

XX

XX

XX

XX

XX

XX

SQ Sequence 564 BP; 137 A; 149 C; 139 G; 139 T; 0 other;

Query Match 79.0%; Score 15.8; DB 22; Length 564;  
Best Local Similarity 89.5%; Pred. No. 2.1e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 ccagttctcaagaagcag 19  
11111111111111111111  
DB 172 ccagttctcaagaagcag 190

## RESULT 14

AA108987  
ID AA108987 standard; DNA: 564 BP.

AC AA108987;

DT 09-OCT-2001 (first entry)

XX Probe #8978 used to measure gene expression in human breast sample.

DE Probe: human; breast disease; breast cancer; development disorder; ss;

KW inflammatory disease; proliferative breast disease; non-carcinoma tumour.

KW Homo sapiens.

PN WO200157270-A2.

PD 09-AUG-2001.

PF 29-JAN-2001; 2001WO-US00661.

PS 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

SQ Sequence 564 BP; 137 A; 149 C; 139 G; 139 T; 0 other;

Query Match 79.0%; Score 15.8; DB 22; Length 564;  
Best Local Similarity 89.5%; Pred. No. 2.1e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 ccagttctcaagagcag 19  
||||||| ||||| |||||  
DB 172 ccagttcacaaagtcag 190

RESULT 15  
AAH00627  
ID AAH00627 standard: DNA; 673.BP.  
AC AAH00627;  
XX  
XX 24-JUL-2001 (first entry)  
DE Enterococcus avium nucleotide sequence SEQ ID NO:618.  
XX  
XX Species specific; genus specific; family specific; probe; detection;  
KW identification; algal; archaeal; bacterial; fungal; parasitica;  
KW microorganism; diagnosis; translation elongation factor Tu; toxin;  
KW translation elongation factor G; RecA recombinase; resistance;  
catalytic subunit of proton-translocating ATPase; antimicrobial;  
vaccine; primer; ds.  
XA  
OS Enterococcus avium.  
XX  
PN WO200123604-A2.  
XX  
XX 05-APR-2001.  
PD  
XX 28-SEP-2000; 2000MO-CA01150.  
PF  
XX 28-SEP-1999; 99CA-2283458.  
PR 19-MAY-2000; 2000CA-2307010.  
XX  
XX (INFE-) INFECTION DIAGNOSTIC (IDI) INC.  
PA Bergeron MG, Boissinot M, Huletsky A, Menard C, Ouellette M;  
PI Picard FJ, Roy PH;  
PI  
XX WPI: 2001-245006/25.  
DR  
XX  
XX Nucleic acid sequences are used to generate universal probes and  
PT primers which can be used to identify and detect the presence of algal,  
PT archaeal, bacterial, fungal and parasitica species in a test sample -  
XX  
XX  
PS Claim 24: Page 763; 1580pp; English.  
CC  
CC The present invention describes a method for generating a repertory of  
nucleic acids of tuf, fus, atpD and/or recA genes from which probes  
and/or primers are derived. The method comprises amplifying the nucleic  
acids of determined algal, archaeal, bacterial, fungal and parasitica  
species with a combination of defined primer pairs. The method can be  
used for producing probes and/or primers for detecting one or more  
related microorganisms e.g. algae, archaea, bacteria, fungi and  
parasites, for universal detection and for specific and ubiquitous  
detection and identification of an algal, archaeal, bacterial, fungal  
and parasitica species, genus, family and group. A nucleic acid (1)  
obtained using the method of the invention can be used for the universal  
detection of any bacterium, fungus or parasite in a sample and for the  
detection of at least one antimicrobial agent resistance gene or at  
least one toxin gene. hexA nucleic acids are used for the specific and  
ubiquitous detection and for identification of Streptococcus pneumoniae.  
(1) can be used to design a therapeutic agent which is effective against  
microorganisms. Microbial species or genus or family or phylum or group  
which can be detected include Abiotrophia adiacens, Bordetella sp.,  
Corynebacterium sp., Enterobacteriaceae group, Escherichia coli,  
Mycobacteriaceae family, Pseudomonads group, Streptococcus sp.,  
Neisseria gonorrhoeae and Staphylococcus sp.. Using DNA based tests  
provides faster results than substrate specificity tests as results can  
be determined in an hour and improved accuracy is also achieved.  
CC AAH00010 to AAH002304 represent nucleotide sequences and primers/probes  
CC which are given in the exemplification of the present invention.  
XX  
XX  
XX Sequence 673 BP: 210 A: 120 C: 154 G: 189 T: 0 other:

Query Match 79.0%; Score 15.8; DB 22; Length 673;  
Best Local Similarity 89.5%; Pred. No. 2.1e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 ccagttctcaagagcag 19  
||||||| ||||| |||||  
DB 205 ccagttctcaagagtcag 223

Search completed: March 4, 2002, 05:21:23  
Job time: 17512 sec

Mon Mar 4 05:51:52 2002

us-08-956-991-5.rml

Page 1

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 4, 2002, 05:11:29 : Search time 226.99 Seconds  
(without alignments)  
19.955 Million cell updates/sec

Title: US-08-956-991-5  
Perfect score: 20  
Sequence: 1 ccaggtctcaagagcagcagg 20

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues  
(number of hits satisfying chosen parameters: 702406)

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PCITUS.C.MB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/Backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed.  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	15.2	76.0	3417	2	US-08-464-402-1
2	15.2	76.0	3417	4	US-09-054-775C-1
3	15.2	76.0	4337	3	US-09-187-049-1
4	15.2	76.0	8298	5	PCT-US93-03076-1
5	15.2	76.0	12752	2	US-08-459-146-1
6	15.2	76.0	12752	2	US-08-459-065-1
7	14.8	74.0	656	4	US-09-328-111-313
8	14.8	74.0	1416	4	US-09-234-393-26
9	14.8	74.0	1785	4	US-09-234-393-51
10	14.8	74.0	1785	4	US-09-234-393-53
11	14.8	74.0	1977	4	US-09-234-393-23
12	14.8	74.0	2071	3	US-09-023-023-1
13	14.8	74.0	8535	3	US-08-716-351A-1
14	14.8	74.0	920	3	US-09-258-373-2
15	14.4	72.0	1314	3	US-07-663-005A-1
16	14.4	72.0	1875	3	US-09-258-373-21
17	14.4	72.0	2135	3	US-08-581-148C-17
18	14.4	72.0	3648	1	US-08-053-614-1
19	14.4	72.0	3648	1	US-08-316-397B-1
20	14.4	72.0	3648	2	US-09-034-306-1
21	14.4	72.0	3648	2	US-09-259-437-1
22	14.4	72.0	3648	5	PCT-US93-09782-1
23	14.4	72.0	4821	1	US-08-053-614-3
24	14.4	72.0	4821	1	US-08-316-397B-3
25	14.4	72.0	4821	2	US-09-034-306-3
26	14.4	72.0	4821	3	US-09-259-437-3
27	14.4	72.0	4821	3	US-09-259-437-3

28	14.4	72.0	4821	5	PCT-US93-09782-3	Sequence 3, Appl1
29	14.4	72.0	5925	3	US-08-470-260-4	Sequence 4, Appl1
30	14.4	72.0	5925	3	US-08-471-491-4	Sequence 4, Appl1
31	14.4	72.0	5925	3	US-08-466-662-4	Sequence 1, Appl1
32	14.4	72.0	10289	2	US-08-477-451-1	Sequence 5, Appl1
33	14.4	72.0	10289	2	US-08-477-451-5	Sequence 25, Appl1
34	14.4	72.0	19932	2	US-08-477-451-25	Sequence 40, Appl1
35	14.2	71.0	401	4	US-09-221-298-40	Sequence 19, Appl1
36	14.2	71.0	460	2	US-08-487-727A-1	Sequence 19, Appl1
37	14.2	71.0	580	2	US-08-272-255-19	Sequence 697, App
38	14.2	71.0	580	5	PCT-US93-08565-19	Sequence 410, App
39	14.2	71.0	688	4	US-08-998-416-697	Sequence 2, Appl1
40	14.2	71.0	870	4	US-09-461-697-410	Sequence 407, App
41	14.2	71.0	874	2	US-08-850-119-2	Sequence 396, App
42	14.2	71.0	911	1	US-08-745-603-1	Sequence 394, App
43	14.2	71.0	936	4	US-09-461-697-402	
44	14.2	71.0	1017	4	US-09-461-697-396	
45	14.2	71.0	1128	4	US-09-461-697-394	

#### ALIGNMENTS

RESULT 1  
US-08-464-402-1  
Sequence 1, Application US/08464402  
Patent No. 5838705  
GENERAL INFORMATION:  
APPLICANT: WEI, ET AL.  
TITLE OF INVENTION: Human DNA Ligase III  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CARELLA, BYRNE, BAIN, GILLFILLAN,  
ADDRESS: CECCHI, STEWART & OLSTEIN  
CITY: ROSELAND  
STREET: 6 BECKER FARM ROAD  
STATE: NEW JERSEY  
COUNTRY: USA  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 INCH DISKETTE  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WORD PERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/464,402  
FILING DATE: June 5, 1995  
CLASSIFICATION: 4.35  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/01939  
FILING DATE: 41 MAR 95  
ATTORNEY/AGENT INFORMATION:  
NAME: FERRARO, GREGORY D.  
REGISTRATION NUMBER: 36,134  
REFERENCE/DOCKET NUMBER: 325800-388  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-994-1740  
TELEFAX: 201-994-1744  
INFORMATION FOR SEQ. ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3417 BASE PAIRS  
TYPE: NUCLEIC ACID  
STRANDEDNESS: SINGLE  
TOPOLOGY: LINEAR  
MOLECULE TYPE: CDNA  
US-08-464-402-1

Query Match 76.0% Score 15.2; DB 2; Length 3417;  
Best Local Similarity 85.0%; Pred. No. 71;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Db 2559 CCACTTGTCCAGAGAGG 2578

## RESULT 2

US-09-054-775C-1  
Sequence 1, Application US/09054775C  
Patent No. 6284504  
GENERAL INFORMATION:  
APPLICANT: Wei, Ying-Fel  
Yu, Guo-Liang  
Haseltine, William  
TITLE OF INVENTION: Human DNA Ligase 111  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: MD  
COUNTRY: USA  
ZIP: 20850

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/054,775C  
FILING DATE: 03-Apr-1998  
CLASSIFICATION: <Unknown>

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/464,402  
FILING DATE: 05-JUN-1995  
APPLICATION NUMBER: PCT/US95/03939  
FILING DATE: 31-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Hoover, Kenley K.  
REGISTRATION NUMBER: 40,302  
REFERENCE/DOCKET NUMBER: P1161D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 301-309-8504  
TELEFAX: 301-309-8439

## INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:  
LENGTH: 3417 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-054-775C-1

Query Match 76.0%; Score 15.2; DB 4; Length 3417;  
Best Local Similarity 85.0%; Pred. No. 74;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ccaagttctcaagagcagg 20  
Db 2559 CCACTTGTCCAGAGAGG 2578

## RESULT 3

US-09-187-049-1  
Sequence 1, Application US/09187049  
Patent No. 6117666  
GENERAL INFORMATION:  
APPLICANT: Lampka, Gayle K.  
TITLE OF INVENTION: PLASTID PROTEOLYTIC PROCESSING ENZYME  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: BRINKS HOFER GILSON & LIONE

STREET: P.O. Box 10395  
CITY: Chicago  
STATE: IL  
COUNTRY: USA  
ZIP: 60610

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/187,049  
FILING DATE:

## CLASSIFICATION:

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/695,177  
FILING DATE:

## ATTORNEY/AGENT INFORMATION:

NAME: Martin, Alice O.  
REGISTRATION NUMBER: 35,601  
REFERENCE/DOCKET NUMBER: 7814/16  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312 321-4200  
TELEFAX: 312 321-4299

## INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:  
LENGTH: 4337 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: unknown  
MOLECULE TYPE: cDNA  
US-09-187-049-1

Query Match 76.0%; Score 15.2; DB 3; Length 4337;  
Best Local Similarity 85.0%; Pred. No. 74;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ccaagttctcaagagcagg 20  
Db 2567 CCACTTGTCCAGAGAGG 2586

## RESULT 4

PCT-US93-03076-1  
Sequence 1, Application PC/TUS9303076  
GENERAL INFORMATION:  
APPLICANT: Whitehead Institute for Biomedical Research  
TITLE OF INVENTION: GAV-Associated protein p190 and  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
STREET: 2 Millia Drive  
CITY: Lexington  
STATE: MA  
COUNTRY: US  
ZIP: 02173  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/03076  
FILING DATE: 19930331  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Granahan, Patricia  
REGISTRATION NUMBER: 32,227  
REFERENCE/DOCKET NUMBER: WHI92-03A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-861-6240

TELEFAX: 617-861-9540  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 8298 base pairs  
 TYPE: NUCLEIC ACID  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 731..5272  
 PCT-US93-03076-1

Query Match 76.0%; Score 15.2; DB 5; Length 8298;  
 Best Local Similarity 85.0%; Pred. No. 82;  
 Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 ccagttctcaaggagcag 20  
 11111111111111111111  
 3763 CCAAGTTCCTCAATGAGACGCTGG 3782

RESULT 5

US-08-459-146-1  
 Sequence 1, Application US/08459146  
 Patent No. 5868405  
 GENERAL INFORMATION:  
 APPLICANT: Choi, Gil Ho  
 APPLICANT: Nuss, Donald Lee  
 TITLE OF INVENTION: Genetically Engineered Transmissible  
 TITLE OF INVENTION: Hypovirulence  
 NUMBER OF SEQUENCES: 3  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: George M. Gould, Esq., Hoffmann-La Roche Inc.  
 STREET: 340 Kingsland Street  
 CITY: Nutley  
 STATE: New Jersey  
 COUNTRY: U.S.A.  
 ZIP: 07110  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/459,146  
 FILING DATE: 02-JUN-1995  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/832,117  
 FILING DATE: 06-FEB-1992  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Roseman, Catherine R  
 REGISTRATION NUMBER: 34,240  
 REFERENCE/DOCKET NUMBER: 8589  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (201) 235-6208  
 TELEFAX: (201) 235-3500  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 12752 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 ORIGINAL SOURCE:  
 ORGANISM: Endothia parasitica (Cryptomycetaria)  
 ORGANISM: parasitica  
 STRAIN: EP713  
 US-08-459-146-1

Query Match 76.0%; Score 15.2; DB 2; Length 12752;  
 Best Local Similarity 85.0%; Pred. No. 88;  
 Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 ccagttctcaaggagcag 20  
 11111111111111111111  
 Db 12192 CCAAGTTCCTCAATGAGACGACG 12211

RESULT 6

US-08-459-065-1  
 Sequence 1, Application US/08459065  
 Patent No. 5882642  
 GENERAL INFORMATION:  
 APPLICANT: Choi, Gil Ho  
 APPLICANT: Nuss, Donald Lee  
 TITLE OF INVENTION: Genetically Engineered Transmissible  
 TITLE OF INVENTION: Hypovirulence  
 NUMBER OF SEQUENCES: 3  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: George M. Gould, Esq., Hoffmann-La Roche Inc.  
 STREET: 340 Kingsland Street  
 CITY: Nutley  
 STATE: New Jersey  
 COUNTRY: U.S.A.  
 ZIP: 07110  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/459,065  
 FILING DATE: 02-JUN-1995  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/832,117  
 FILING DATE: 06-FEB-1992  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Roseman, Catherine R  
 REGISTRATION NUMBER: 34,240  
 REFERENCE/DOCKET NUMBER: 8589  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (201) 235-6208  
 TELEFAX: (201) 235-3500  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 12752 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 ORIGINAL SOURCE:  
 ORGANISM: Endothia parasitica (Cryptomycetaria)  
 ORGANISM: parasitica  
 STRAIN: EP713  
 US-08-459-065-1

Query Match 76.0%; Score 15.2; DB 2; Length 12752;  
 Best Local Similarity 85.0%; Pred. No. 88;  
 Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 ccagttctcaaggagcag 20  
 11111111111111111111  
 Db 12192 CCAAGTTCCTCAATGAGACGACG 12211

RESULT 7





OTHER INFORMATION: Description of Artificial Sequence:  
 OTHER INFORMATION: computer-generated nucleic acid sequence encoding  
 OTHER INFORMATION: gamma humulene synthase  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: (4)..(1782)  
 US-09-234-393-51

Query Match 74.0%; Score 14.8; DB 4; Length 1785;  
 Best Local Similarity 88.9%; Pred. No. 1e+02;  
 Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 agtctcaaaagagcagc 20  
 ||||| |||| |||||  
 Db 1616 agtctcaaaagagcagc 1633

RESULT 11  
 US-234-393-53  
 Sequence 53, Application US/09234393A  
 Patent No. 6265639

GENERAL INFORMATION:  
 APPLICANT: Croteau, Rodney B  
 APPLICANT: Croteau, John E  
 APPLICANT: Bohman, Jorg  
 APPLICANT: Jetter, Reinhard  
 APPLICANT: Steele, Christopher L  
 TITLE OF INVENTION: SESOUTEREPNE SYNTHASES FROM GRAND FIR (ABIES GRANDIS)  
 FILE REFERENCE: MSUR113345  
 CURRENT APPLICATION NUMBER: US/09/234,393A  
 EARLIER FILING DATE: 1999-01-20  
 EARLIER FILING DATE: 1998-01-22  
 NUMBER OF SEQ ID NOS: 55  
 SOFTWARE: Patentin Ver. 2.0  
 SEQ ID NO 53  
 LENGTH: 1785  
 TYPE: DNA  
 ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: Description of Artificial Sequence:  
 OTHER INFORMATION: computer-generated nucleic acid sequence encoding  
 OTHER INFORMATION: gamma humulene synthase  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: (4)..(1782)  
 US-234-393-53

Query Match 74.0%; Score 14.8; DB 4; Length 1785;  
 Best Local Similarity 88.9%; Pred. No. 1e+02;  
 Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 agtctcaaaagagcagc 20  
 ||||| |||| |||||  
 Db 1616 agtctcaaaagagcagc 1633

RESULT 12  
 US-09-234-393-23  
 Sequence 23, Application US/09234393A  
 Patent No. 6265639

GENERAL INFORMATION:  
 APPLICANT: Croteau, Rodney B  
 APPLICANT: Croteau, John E  
 APPLICANT: Bohman, Jorg  
 APPLICANT: Jetter, Reinhard  
 APPLICANT: Steele, Christopher L  
 TITLE OF INVENTION: SESOUTEREPNE SYNTHASES FROM GRAND FIR (ABIES GRANDIS)  
 FILE REFERENCE: MSUR113345

CURRENT APPLICATION NUMBER: US/09/234,393A  
 CURRENT FILING DATE: 1999-01-20  
 EARLIER APPLICATION NUMBER: 60/072,204  
 EARLIER FILING DATE: 1998-01-22  
 NUMBER OF SEQ ID NOS: 55  
 SOFTWARE: Patentin Ver. 2.0  
 SEQ ID NO 23  
 LENGTH: 1977  
 TYPE: DNA  
 ORGANISM: Abies grandis  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: (4)..(1782)  
 US-09-234-393-24

Query Match 74.0%; Score 14.8; DB 4; Length 1977;  
 Best Local Similarity 88.9%; Pred. No. 1e+02;  
 Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 agtctcaaaagagcagc 20  
 ||||| |||| |||||  
 Db 1616 agtctcaaaagagcagc 1633

RESULT 13  
 US-09-023-023-1/c  
 Sequence 1, Application US/09023023  
 Patent No. 6121018

GENERAL INFORMATION:  
 APPLICANT: Kristine Kay Kikly  
 TITLE OF INVENTION: Interleukin-1 Beta Converting Enzyme Like Apoptosis Pr  
 NUMBER OF SEQUENCES: 2  
 CORRESPONDENCE ADDRESS:  
 ADDRESS: Smithkline Beecham Corporation  
 STREET: 709 Swedeland Road  
 CITY: King Of Prussia  
 STATE: PA  
 COUNTRY: USA  
 ZIP: 19406  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Disket  
 COMPUTER: IBM Compati  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FastSeq for Windows Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/023,023  
 FILING DATE: 12-FEB-1998  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 60/042,010  
 FILING DATE: March 27, 1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Han, William T  
 REGISTRATION NUMBER: 34,344  
 REFERENCE/DOCKET NUMBER: GH50013  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 610-270-5219  
 TELEFAX: 610-270-5090  
 TELEX:  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2071 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 US-09-023-023-1

Query Match 74.0%; Score 14.8; DB 3; Length 2071;  
 Best Local Similarity 88.9%; Pred. No. 1e+02;  
 Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 agtctcaagagcagcag 20  
 ||| |||||  
 DB 1545 AGTGATCAAGCAGCAGC 1528

RESULT 14

US-08-716-351A-1

Sequence 1, Application US/08716351A  
 Patent No. 6033905

GENERAL INFORMATION:

APPLICANT:

TITLE OF INVENTION: Gibbon Ape Leukemia Virus-Based

TITLE OF INVENTION: Retroviral Vectors

NUMBER OF SEQUENCES: 5

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/716,351A

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US94/03784

FILING DATE: 06-APR-1994

ATTORNEY/AGENT INFORMATION:

NAME: Bastian, Kevin L.

REGISTRATION NUMBER: 34,774

REFERENCE/DOCKET NUMBER: 15280-128-1PC

TELEPHONE: (415) 543-9600

TELEFAX: (415) 543-5043

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 8535 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

FEATURE:

NAME/KEY: misc.feature

LOCATION: 1..8535

OTHER INFORMATION: /standard\_name="Galy SATO Genome"

US-08-716-351A-1

Query Match 74.0%; Score 14.8; DB 3; Length 8535;  
 Best Local Similarity 88.9%; Pred. No. 1.3e+02;  
 Matches 16; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 2 cagtctcaagagcagcag 19

||| |||||  
 DB 1977 CAGGTCTCAAGCAGCAG 1994

RESULT 15

US-09-258-373-2/c

Sequence 2, Application US/09258373

Patent No. 6150110

GENERAL INFORMATION:

APPLICANT: Fletcher, Jonathan A.

TITLE OF INVENTION: HMGI(Y)-LAMA4\* FUSION ONCOGENE,

TITLE OF INVENTION: ONCOPROTEIN AND METHODS OF USE

FILE REFERENCE: B0801/7135/ERP

CURRENT APPLICATION NUMBER: US/09/258,373

CURRENT FILING DATE: 1999-02-26

EARLIER APPLICATION NUMBER: 60/076,401

EARLIER FILING DATE: 1998-02-26

NUMBER OF SEQ ID NOS: 22

SOFTWARE: FASTSEQ for Windows Version 3.0

; SEQ ID NO 2  
 ; LENGTH: 920  
 ; TYPE: DNA  
 ; ORGANISM: Homo Sapiens  
 US-09-258-373-2

Query Match 72.0%; Score 14.4; DB 3; Length 920;  
 Best Local Similarity 93.8%; Pred. No. 1.4e+02;  
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 5 ttctcaagagcagcag 20  
 |||||  
 DB 225 TTCTCAAGCAGCAGC 210

Search completed: March 4, 2002, 05:11:33  
 Job time: 16947 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 4, 2002, 02:14:35 ; Search time 6206.58 Seconds  
(without alignments)  
34.627 Million cell updates/sec

Title: US-08-956-991-6

Perfect score: 20

Sequence: 1 cctgatgacctgcaggaag 20

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1151937 segs, 537289281 residues  
22703874

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*  
1: em\_estfun:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estom:\*  
5: em\_estpl:\*  
6: em\_estda:\*  
7: em\_estro:\*  
8: em\_estov:\*  
9: em\_hnc:\*  
10: gb\_estl:\*  
11: gb\_est2:\*  
12: gb\_hnc:\*  
13: gb\_gss:\*  
14: em\_gss\_fun:\*  
15: em\_gss\_hum:\*  
16: em\_gss\_hiv:\*  
17: em\_gss\_pla:\*  
18: em\_gss\_pro:\*  
19: em\_gss\_rtd:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	744	11	BF348620 602019924
2	18.4	92.0	424	13	AZ297511 KFC1-23-1
3	17.4	87.0	903	13	CNS05070 AL315219 Tetradon
4	17	85.0	559	11	BG700618 BG700618
5	17	85.0	655	11	BG701055 602682072
6	16.8	84.0	176	10	BB373590 BB373590
7	16.8	84.0	188	11	F25472 HSPD12539 H
8	16.8	84.0	270	10	BB528026 BB528026
9	16.8	84.0	274	10	BB367406 BB367406
10	16.8	84.0	276	10	BE032419 171758 MA
11	16.8	84.0	278	10	BB604281 BB604281
12	16.8	84.0	291	10	BB429902 BB429902

13	16.8	84.0	292	11	BF825905
14	16.8	84.0	298	10	BB116563
15	16.8	84.0	334	10	AI841727
16	16.8	84.0	338	11	F00929
17	16.8	84.0	339	10	AA855923
18	16.8	84.0	349	10	AI836340
19	16.8	84.0	349	11	R05334
20	16.8	84.0	350	10	AI905927
21	16.8	84.0	358	11	BF651521
22	16.8	84.0	359	11	BF825966
23	16.8	84.0	368	11	BF826657
24	16.8	84.0	373	11	BE899939
25	16.8	84.0	381	10	AV592640
26	16.8	84.0	388	10	AJ272002
27	16.8	84.0	401	11	AJ278473
28	16.8	84.0	401	11	BF738908
29	16.8	84.0	416	10	AA65681
30	16.8	84.0	428	11	BF826523
31	16.8	84.0	445	11	BF795797
32	16.8	84.0	448	11	BF873693
33	16.8	84.0	450	11	BF826402
34	16.8	84.0	459	11	BG610562
35	16.8	84.0	459	10	BE233426
36	16.8	84.0	460	10	AJ278546
37	16.8	84.0	466	13	AW535840
38	16.8	84.0	467	13	AZ851563
39	16.8	84.0	468	11	BF826596
40	16.8	84.0	471	11	BG796032
41	16.8	84.0	478	11	BF592276
42	16.8	84.0	485	11	BF599809
43	16.8	84.0	485	11	AO815436
44	16.8	84.0	487	11	BF603509
45	16.8	84.0	490	11	BG794127

## ALIGNMENTS

RESULT 1  
LOCUS BF348620 744 bp mRNA  
DEFINITION 602019924F1 NCI CGAP\_Brn67 Homo sapiens cDNA clone IMAGE:4155508  
5', mRNA sequence.  
ACCESSION BF348620  
VERSION BF348620.1 GI:11296215  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
REFERENCE 1 (bases 1 to 744)  
AUTHORS NIH-MGC http://mgc.ncl.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-femail.nih.gov  
Tissue: Procurement: David N. Louis, M.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: L1AM426 row: p column: 05  
High quality sequence stop: 697.  
Location/Qualifiers  
1. 744  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4155508"  
/clone\_1ib="NCI-CGAP\_Brn67"  
/rname="type="anaplastic oligodendroglioma with 1p/19q loss"

## FEATURES

source



ACCESSION BG700618  
 VERSION BG700618.1 GI:13970140  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 REFERENCE: NIH-MGC <http://mgi.nci.nih.gov/>.  
 (bases 1 to 559)  
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
 TITLE Unpublished (1999)  
 JOURNAL Contact: Robert Strausberg, Ph.D.  
 COMMENT Email: [cgabs@mail.nih.gov](mailto:cgabs@mail.nih.gov)  
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
 cDNA Library Preparation: Michael J. Brownstein (NMGRI), Shitaki  
 Toshitsuki and Piero Carninci (RIKEN)  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium <http://llnl.ac>:  
<http://image.llnl.gov>  
 Plate: LLAM10711 row: m column: 03  
 High quality sequence stop: 559.  
 Location/Qualifiers

FEATURES  
 source  
 1..559  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_image="4815146"  
 /clone\_lib="NIH-MGC\_95"  
 /issue\_type="hippocampus"  
 /lab\_host="DH10B"  
 /note="organ: brain; Vector: pBluescript (modified  
 pBluescript KS+); Site\_1: BamHI; Site\_2: SalI-XhoI (gtcagag  
 ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTNN-3',  
 size-selected for average insert size 2.5 kb and  
 normalized to KOF 5. This is a primary library enriched  
 for full-length clones and constructed using the  
 cap-trapper method (Carninci, in preparation). Library  
 constructed by M. Brownstein (NMGRI), National  
 Institutes of Health). Note: this is a NIH-MGC Library."

HASH COUNT 175 a 114 c 139 g 131 t  
 ORIGIN

Query Match 85.0%; Score 17; DB 11; Length 559;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+03;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 cctglatgacctgcag 17  
 |||||||  
 121 CCTGTATGACCTGCAG 105

RESULT 5  
 BG701055 655 bp mRNA EST 07-MAY-2001  
 602662072p1 NIH-MGC\_95 Homo sapiens cDNA clone IMAGE:4814849 5',  
 mRNA sequence.  
 ACCSSION BG701055  
 VERSION BG701055.1 GI:13971014  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE: NIH-MGC <http://mgi.nci.nih.gov/>.  
 (bases 1 to 655)  
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
 TITLE Unpublished (1999)  
 JOURNAL Contact: Robert Strausberg, Ph.D.  
 COMMENT Email: [cgabs@mail.nih.gov](mailto:cgabs@mail.nih.gov)  
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
 cDNA Library Preparation: Michael J. Brownstein (NMGRI), Shitaki

Toshitsuki and Piero Carninci (RIKEN)  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium <http://llnl.ac>:  
<http://image.llnl.gov>  
 Plate: LLAM10710 row: p column: 18  
 High quality sequence stop: 620.  
 Location/Qualifiers

FEATURES  
 source  
 1..655  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_image="4814849"  
 /clone\_lib="NIH-MGC\_95"  
 /issue\_type="hippocampus"  
 /lab\_host="DH10B"  
 /note="organ: brain; Vector: pBluescript (modified  
 pBluescript KS+); Site\_1: BamHI; Site\_2: SalI-XhoI (gtcagag  
 ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTNN-3',  
 size-selected for average insert size 2.5 kb and  
 normalized to KOF 5. This is a primary library enriched  
 for full-length clones and constructed using the  
 cap-trapper method (Carninci, in preparation). Library  
 constructed by M. Brownstein (NMGRI), National  
 Institutes of Health). Note: this is a NIH-MGC Library."  
 BASE COUNT 205 a 140 c 154 g 156 t  
 ORIGIN

Query Match 85.0%; Score 17; DB 11; Length 655;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+03;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cctglatgacctgcag 17  
 |||||||  
 Db 121 CCTGTATGACCTGCAG 105

RESULT 6  
 BB373590 EST 13-JUL-2000  
 LOCUS BB373590  
 DEFINITION BB373590 KIKEN full-length enriched, 16 days embryo head Mus  
 musculus cDNA clone G130071C07 3' similar to AB025261 Sus scrofa  
 mRNA for myosin heavy chain 2b, mRNA sequence.  
 BB373590.1 GI:1086084  
 EST.  
 VERSION BB373590.1 GI:1086084  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE: Kono, H., Aizawa, K., Akahira, S., Akiyama, J., Arikawa, T., Carninci  
 1 (bases 1 to 176)  
 AUTHORS Kono, H., Aizawa, K., Akahira, S., Akiyama, J., Arikawa, T., Carninci  
 P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N.,  
 Hirozane, T., Hori, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M.,  
 Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N.,  
 Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C., Kusabe, M.,  
 Matsuyama, T., Miki, K., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y.,  
 Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K., Shibata, K., Shibata  
 Y., Shigemoto, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Sugahara, Y.,  
 Suzuki, H., Suzuki, H., Tagawa, A., Takahashi, F., Tomimaga, N., Toya  
 T., Tsunoda, Y., Watahiki, A., Watanabe, S., Yamamura, T., Yananaka, I.,  
 Yano, K., Yasunishi, A., Yokota, T., Yoshida, K., Yoshiki, A., Yoshino  
 M., Muramatsu, M. and Hayashizaki, Y.  
 RIKEN Mouse ESTs (Kono, H., et al.)  
 Unpublished (2000)  
 TITLE Contact: Yoshinori Hayashizaki  
 JOURNAL Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Saitama-shi, Saitama-shi, Saitama-shi, Saitama-shi, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216



The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Sphiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@sc.riken.go.jp,  
url: http://genome.riken.go.jp/  
Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagasaka, S., Saito, N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Thermostabilization and thermoactivation of thermolabile enzymes for trehalose and its application for the synthesis of full length cDNA. *Proc Natl. Acad. Sci. U.S.A.* 95 (2), 520-524 (1998)  
Itoh, M., Kitsuai, T., Akiyama, J., Shibata, K., Izawa, M., Kawaji, Y., Tomaru, Y., Carninci, P., Shibata, T., Ozawa, Y., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.  
Automated filtration-based high-throughput plasmid preparation system. *Genome Res.* 9 (5), 463-470 (1999)  
Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning. *Methods Enzymol.* 303, 19-44 (1999)  
Please visit our web site (<http://genome.riken.go.jp/>) for further details.

Location/Qualifiers  
1. .270

FEATURES

SOURCE

Email: genome-res@sc.riken.go.jp,  
URL: <http://genome.riken.go.jp/>  
Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoaka, S., Sasaki,  
N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.,  
Thermotabilization and thermooactivation of thermostable enzymes by  
trehalose and its application for the synthesis of full length  
cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)  
Itoh, M., Kishimui, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J.,  
Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki,  
Y. and Hayashizaki, Y.,  
Automated filtration-based high-throughput plasmid preparation  
system. Genome Res. 9 (5), 463-470 (1999)  
Carninci, P. and Hayashizaki, Y.,  
High-efficiency full-length cDNA cloning. Methods Enzymol. 303,  
19-44 (1999)  
Please visit our web site (<http://genome.riken.go.jp/>) for  
further details.

Location/Qualifiers

1. 274

/organism="Mus musculus"  
/sitein="5,7H,6J"  
/db\_xref="taxon:10090"  
/clone="C140037M06"  
/contig\_id="RIKEN full-length enriched, 16 days embryo  
head"  
/sex="male"  
/issue\_type="head"  
/dev\_stage="16 days embryo"

/M01c-Site1: salt, site2: samir, ob-  
prepared and sequenced in Mouse Genome Encyclopedia  
Project of Genome Exploration Research group in Riken  
Genomic Sciences Center and Genome Science Laboratory  
in RIKEN. Division of Experimental Animal Research in Riken  
contributed to prepare mouse tissues. 1st strand cDNA was  
primed with a primer [5',  
GAGAGGAAAGATCCAGAGCTCTTTTTTTTTTTTNN 3'], cDNA was

transcription factor c-Jun. Second strand cDNA was prepared with the primer dAapler of sequence [5', GAGAGAGCAATTCGAGCTAATTAAATTAATGCCCCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from LambdaFLC<sup>1a</sup>.

	ORGANISM	Mus musculus Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. (Class 1 to 278)
	REFERENCE	Alzawa,K., Akhira,S., Akimura,T., Arai,A., Arakawa,T., Carninci,P., Hanaagaki,T., Hayatsu,N., Hirooka,T., Hirozane,T., Hodyama,Y., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kawai,J., Koijima,Y., Konno, H., Kusakabe,M., Matsuyama,T., Miyazaki,A., Nakamura,M., Nishi,K., Nomura,K., Nunazaki,R., Okazaki,Y., Okido,T., Owa,C., Sakai,C., Sakai,K., Sasaki,D., Salo,K., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Tanaka, T., Toyota,T., Wachihiiki,A., Yamamura,T., Yasunishi,A., Yoshida,K., Yoshiki,A., Muramatsu,M. and Hayashizaki,Y.
	JOURNAL	RIKEN Mouse Esrs (Alzawa,K. et al. 2000) Unpublished (2000)
	COMMENT	Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSCC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel.: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-resgsc.riken.go.jp, urc:http://genome-qsc.riken.go.jp/ Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S., Sasaki, N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y. Thermostabilization and thermoinactivation of thermostable enzymes by trehalose and its application for the syntheses of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2): 520-524 (1998) Itoh,M., Kitsumori,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J., Toraru,Y., Carninci,P., Shibata,Y., Ozawa,T., Muramatsu,M., Okazaki, Y. and Hayashizaki,Y. Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5): 463-470 (1999) Ganninci,P. and Hayashizaki,Y. High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999) Please visit our web site ( <a href="http://genome.rtc.riken.go.jp">http://genome.rtc.riken.go.jp</a> ) for further details.
FEATURES	SOURCE	location/qualifiers 1..278 /organism "Mus musculus" /strain "57RU/61" /db_xref "taxon:10090" /clone "p140014615" /contig_1ib "RIKEN full length enriched, 15 days embryo head" /sex "mixed" /library_type "head" /dev_stage "15 days embryo" /lab_host "DH10B" /note "Site 1: Salt. Site 2: BamHI. cDNA library was prepared and sequenced in Mouse Genome encyclopedia Project at Genome Exploration Research group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken continued to prepare mouse tissues. 1st strand cDNA was primed with a primer 15'. GGAGGACGAGACATCCACACGCTTTTTTTTTTTTTVN 3'}. cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence 15'. GGAGGACGATTTCGCAGCTTAATTAATTAATGCCCCCCCCC 3'}. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from Lambda phage 1*
BASE COUNT	79 a 53 C 98 G 48 t	
ORIGIN		
Query Match	84.0%	Score 16.8; DB 10; Length 278;
Best Local Similarity	90.0%;	Pred. No. 1.7e+03;







```

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UI-M-AL0-abo-f-04-0-UI"
/clone_lib="NIH_BMAP_MCO"
/dev_stage="27-32 days"
/lab_host="DH10B (Life Technologies)"
/notes="Vector: p7773D-Pac (Pharmacia) with a modified
polylinker: Site_1: Not I; Site_2: Eco RI; The
NIH_BMAP_MCO library is a non-normalized library
constructed from mouse cortex. The tag is a string of 5
nucleotides present between the Not I site and the
oligo-dt track. The library was constructed as described
by Bonaldo, Lennon and Soares, Genome Research 6: 791-806
, 1996. Tissue provided by Ms. Annie Novakovich,
Zivic-Miller Laboratories.
TAG_LIB=NIH_BMAP_MCO
TAG_TISSUE=prefrontal-cortex
TAG_SEQ=GCACA"
BASE COUNT      60 a      92 c      100 g      82 t
      TN

```

```

Query Match      84.0%; Score 16.8; Dh 10; Length 334.
Best Local Similarity 90.0%; Pred. No. 1.7e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Caps 0;
QY      1 cctgtagacctgcaggaag 20
      ||||| ||||| ||||| |||||
DB      215 CCTGTGTGACCTCGCAGGATG 196

```

Search completed: March 4, 2002, 02:14:39  
 Job time: 6488 sec



developmental and neurological abnormalities

XX  
XX  
PS Example 5: Page 79: 109pp: English.

CC Reverse primer B9-131R and forward primer B9-131F (see AAV31983)

CC were used in RT-PCR assays of cDNA libraries of various human

CC tissues to determine human Down syndrome-cell adhesion molecule

CC DS-CAM (see AAV31981 and AAV31988) expression. The results

CC demonstrated expression of human DS-CAM mRNA in foetal and adult

CC brain, and foetal kidney. In addition, a breast carcinoma cell

CC line showed expression of human DS-CAM mRNA. DS-CAM polypeptides

CC (see AAM42086-87) are associated with developmental and neurological

CC processes. The polypeptides and nucleic acids are used to develop

CC products for the detection, diagnosis and therapy of developmental

CC and neurological abnormalities.

CC  
CC  
SQ Sequence 20 BP: 5 A: 5 C: 6 G: 4 T: 0 other:

TRY Match 100.0%; Score 20; DI 19; Length 20;

at Local Similarity 100.0%; Pred. No. 4.5;

Matches 20: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 cctgatagacctgcaggaag 20  
1 cctgatagacctgcaggaag 20

Db 1 cctgatagacctgcaggaag 20

RESULT 2

AAV31982/C

ID AAV31982 standard: cDNA: 388 BP.

XX  
XX  
AC AAV31982:

XX  
XX  
DT 28-SEP-1998 (first entry)

XX  
XX  
DE Human Down syndrome-cell adhesion molecule DS-CAM probe E51.

XX  
XX  
DS-CAM1: Down syndrome-cell adhesion molecule: neural cell;

KM signal transduction: trisomy 21: mental retardation;

KM holoprosencephaly: corpus callosum agenesis;

KM schizencephaly: diagnosis: assay: human: ds: ss.

XX  
XX  
OS Homo sapiens.

XX  
XX  
PN W09817795-A1.

XX  
XX  
P 30-APR-1998.

XX  
XX  
P 23-OCT-1997: 97MO-US19547.

XX  
XX  
PR 25-OCT-1996: 96US-0029322.

XX  
XX  
PA (CEDA-) CEDARS SINAI MEDICAL CENT.

XX  
XX  
PI Korenberg JK:

XX  
XX  
P1 WPI: 1998-271791/24.

XX  
XX  
DR WPI: 1998-271791/24.

XX  
XX  
PT New isolated Down's Syndrome-cell adhesion molecule - used to

PT develop products for detection, diagnosis and therapy of

PT developmental and neurological abnormalities

XX  
XX  
PS Example 2: Page 79: 109pp: English.

XX  
XX  
CC This polynucleotide comprises cDNA fragment E51 that was isolated

CC from a human trisomy 21 foetal brain (14-wk gestation) cDNA

CC library following a modified direct cDNA selection technique

CC applied to bacterial and pl artificial chromosomes between ETS2

CC and MX1. E51 was used as a probe to screen the trisomy 21 foetal

CC brain-library. 62 clones were isolated from the 2 million clones

CC of the original library. Overlapping clones were obtained that

CC were used to deduce a full-length sequence (see AAV31981) coding

CC for novel Down syndrome-cell adhesion molecule DS-CAM1 (see

CC AAM42086). A splice variant, DS-CAM2 (see AAV31988), was also

CC identified. DS-CAM polypeptides are associated with developmental

CC and neurological processes. The polypeptides and nucleic acids

CC are used to develop products for the detection, diagnosis and

CC therapy of developmental and neurological abnormalities.

XX  
XX  
SQ Sequence 388 BP: 83 A: 119 C: 95 G: 91 T: 0 other:

Query Match 100.0%; Score 20; DI 19; Length 388;

Best Local Similarity 100.0%; Pred. No. 5.3;

Matches 20: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 cctgatagacctgcaggaag 20  
1 cctgatagacctgcaggaag 20

Db 331 CCGTATGACCTGCGCAG 312

RESULT 3

AAV31981

ID AAV31981 standard: cDNA: 6604 BP.

XX  
XX  
AC AAV31981:

XX  
XX  
DT 28-SEP-1998 (first entry)

XX  
XX  
DE Human Down syndrome-cell adhesion molecule DS-CAM1 cDNA.

XX  
XX  
DS-CAM1: Down syndrome-cell adhesion molecule: neural cell;

KM signal transduction: trisomy 21: mental retardation;

KM holoprosencephaly: corpus callosum agenesis;

KM schizencephaly: diagnosis: assay: human: ds: ss.

XX  
XX  
OS Homo sapiens.

XX  
XX  
PN W09817795-A1.

XX  
XX  
P 30-APR-1998.

XX  
XX  
P 23-OCT-1997: 97MO-US19547.

XX  
XX  
PR 25-OCT-1996: 96US-0029322.

XX  
XX  
PA (CEDA-) CEDARS SINAI MEDICAL CENT.

XX  
XX  
PI Korenberg JK:

XX  
XX  
P1 WPI: 1998-271791/24.

XX  
XX  
DR P-PSDB: AAM42086.

XX  
XX  
PT New isolated Down's Syndrome-cell adhesion molecule - used to

PT develop products for detection, diagnosis and therapy of

PT developmental and neurological abnormalities

XX  
XX  
PS Claim 3: Page 65-72: 109pp: English.

XX  
XX  
CC This cDNA clone codes for Down syndrome-cell adhesion molecule

CC DS-CAM1 (see AAM42086), a cell surface glycoprotein belonging to a

CC novel subclass of the Ig superfamily with highest homology to

CC neural cell adhesion molecules. A modified direct cDNA selection

CC technique was applied to bacterial and pl artificial chromosomes

CC between ETS2 and MX1 by using cDNA from trisomy 21 human foetal

CC brain. A unique cDNA fragment, designated E51 (see AAV31982), was

CC obtained and used to screen a trisomy 21 human foetal brain (14-wk

CC gestation) cDNA library. Further clones were obtained by exon  
CC trapping. The complete DS-CAM2 cDNA sequence was deduced from  
CC overlapping clones. A splice variant cDNA (see AAV31988), encoding  
CC non-membrane bound DS-CAM2 (see AAM42087), was also identified. The  
CC DS-CAM gene spans 900-1200 kb of genomic DNA and maps at chromosome  
CC 21q22.2-22.3. The invention also provides murine DS-CAM partial  
CC sequences (see AAV31985-87), expression vectors and primers useful  
CC for amplification of DS-CAM nucleic acid. DS-CAM polypeptides are  
CC associated with developmental and neurological processes. They can  
CC be used in e.g. neural prosthetic devices used in entubulation  
CC methods of repairing (regenerating) damaged or severed peripheral  
CC nerves. The products can also be used in detection, diagnosis and  
CC therapy of developmental and neurological abnormalities such as  
CC Down syndrome, mental retardation, holoprosencephaly, agenesis of  
CC the corpus callosum, or schizencephaly. Antisense oligonucleotides  
CC are used for inhibiting translation of mRNA.

CC Sequence 6604 BP; 1673 A; 1827 G; 1336 T; 0 other;

CC Query Match 100.0%; Score 20; DB 19; Length 6604;  
CC Best Local Similarity 100.0%; Pred. NO. 6.2; Mismatches 0; Gaps 0;  
CC Matches 20; Conservative 0; Indels 0; Caps 0;

OY 1 cctgatagcctgcaggaag 20  
Db 5054 cctgatagcctgcaggaag 5073

# RESULT 4

AAV31988 standard; cDNA: 6413 BP.

AC AAV31988;

DT 28-SEP-1998 (first entry)

XX Human Down syndrome-cell adhesion molecule DS-CAM2 cDNA.

XX DS-CAM2; Down syndrome-cell adhesion molecule; neural cell;

KW signal transduction; trisomy 21; mental retardation;

KW holoprosencephaly; corpus callosum agenesis;

KW schizencephaly; diagnosis; assay; human; ds; ss.

XX Homo sapiens.

XX Key location/Qualifiers

XX CDS 453..5168

XX /\*tag= a

XX MO9817795-A1.

XX 30-APR-1998.

XX 23-OCT-1997; 97WO-0519547.

XX 25-OCT-1996; 96US-0029322.

XX (CEDA-) CEDARS SINAI MEDICAL CENT.

XX Korenberg JR;

XX MPI: 1998-271791/24.

XX P-PSDH; AAM42087.

XX New isolated Down's Syndrome-cell adhesion molecule - used to

XX develop products for detection, diagnosis and therapy of

XX developmental and neurological abnormalities

XX Claim 3: Page 83-90; 107pp; English.

XX This cDNA clone codes for Down syndrome-cell adhesion molecule

CC DS-CAM2 (see AAM42087), an extracellular, soluble protein belonging  
CC to a novel subclass of the Ig superfamily with highest homology to  
CC neural cell adhesion molecules. A modified direct cDNA selection  
CC technique was applied to bacterial and plasmid artificial chromosomes  
CC between ETS2 and Xkl by using cDNA from trisomy 21 human foetal  
CC brain. A unique cDNA fragment, designated E51 (see AAV31982), was  
CC obtained and used to screen a trisomy 21 human foetal brain (14-wk  
CC gestation) cDNA library. Further clones were obtained by exon  
CC trapping, and the complete DS-CAM2 cDNA sequence was identified.  
CC DS-CAM2 is a splice variant of membrane-bound DS-CAM (see AAM42086)  
CC and is deleted for the entire transmembrane domain. The DS-CAM  
CC gene spans 900-1200 kb of genomic DNA and maps at chromosome  
CC 21q22.2-22.3. The invention also provides murine DS-CAM partial  
CC sequences (see AAV31985-87), expression vectors and primers useful  
CC for amplification of DS-CAM nucleic acid. DS-CAM polypeptides are  
CC associated with developmental and neurological processes. They can  
CC be used in e.g. neural prosthetic devices used in entubulation  
CC methods of repairing (regenerating) damaged or severed peripheral  
CC nerves. The products can also be used in detection, diagnosis and  
CC therapy of developmental and neurological abnormalities such as  
CC Down syndrome, mental retardation, holoprosencephaly, agenesis of  
CC the corpus callosum, or schizencephaly. Antisense oligonucleotides  
CC are used for inhibiting translation of mRNA.

CC Sequence 6413 BP; 1633 A; 1781 G; 1707 G; 1292 T; 0 other;

CC Query Match 92.0%; Score 18.4; DB 19; Length 6413;  
CC Best Local Similarity 95.0%; Pred. NO. 32;  
CC Matches 19; Conservative 0; Mismatches 1; Indels 0; Caps 0;

OY 1 cctgatagcctgcaggaag 20  
Db 5054 cctgatagcctgcaggaag 5073

# RESULT 5

AAV31987 standard; cDNA: 2173 BP.

AC AAV31987;

DT 28-SEP-1998 (first entry)

XX Mouse Down syndrome-cell adhesion molecule DS-CAM 3' cDNA.

XX DS-CAM; Down syndrome-cell adhesion molecule; neural cell;

KW signal transduction; trisomy 21; mental retardation;

KW holoprosencephaly; corpus callosum agenesis;

KW schizencephaly; diagnosis; assay; mouse; ds; ss.

XX Mus sp.

XX MO9817795-A1.

XX 30-APR-1998.

XX 23-OCT-1997; 97WO-0519547.

XX 25-OCT-1996; 96US-0029322.

XX (CEDA-) CEDARS SINAI MEDICAL CENT.

XX Korenberg JR;

XX MPI: 1998-271791/24.

XX New isolated Down's Syndrome-cell adhesion molecule - used to

XX develop products for detection, diagnosis and therapy of

XX developmental and neurological abnormalities

XX Claim 2: Page 81-83; 107pp; English.







```

XX
DR WPI: 2001-488897/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human placenta -
XX
PS Claim 25: SEQ ID No 3938; 654pp; English.
XX
XX The present invention relates to single exon nucleic acid probes (SENPs).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders.
XX
SQ Sequence 1954 BP; 634 A; 350 C; 447 G; 523 T; 0 other:

Query Match      84.0%; Score 16.8; DB 22; Length 1954;
Best Local Similarity 90.0%; Pred. No. 1.5e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

{
  0Y      1 cctgtatgacctgcagaag 20
          11111111111111111111
  Db      200 ccagatagacctgcagaag 219

RESULT 10
AA103753
ID AA103753 standard; DNA; 1954 BP.
XX
XX AA103753;
XX
XX 09-OCT-2001 (first entry)
XX
XX Probe #3744 used to measure gene expression in human breast sample.
DE
XX
XX Probe: human; breast disease; breast cancer; development disorder; ss;
KM Inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
XX Homo sapiens.
OS
XX
XX WO200157270-A2.
PN
XX
XX 09-AUG-2001.
PD
XX
XX 29-JAN-2001; 2001WO-US00661.
PF
XX
XX 04-FEB-2000; 2000US-0180312.
H 26-MAY-2000; 2000US-0207456.
H 30-JUN-2000; 2000US-0608408.
PH 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX
XX Penn SG, Hanzel DK, Chen W, Rank DK;
P1
XX
XX WPI: 2001-476286/51.
DR
XX
XX Novel single exon nucleic acid probe used to measuring gene expression
PT in a human breast -
XX
XX Claim 25: SEQ ID No 3744; 322pp; English.
XX
XX The present invention relates to novel single exon nucleic acid probes.
CC The present sequence is one such probe. The probes are useful for
CC measuring human gene expression in a human breast sample, where the probe
CC hybridises at high stringency to a nucleic acid expressed in the human
CC breast. The probes are useful for predicting, diagnosing, grading,
CC staging, monitoring and prognosing diseases of the human breast,
CC particularly those diseases with polygenic aetiology. The diseases

```

```

CC include: breast cancer, disorders of development, inflammatory diseases
CC of the breast, fibrocystic changes, proliferative breast disease and
CC non-carcinoma tumours.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1954 BP; 634 A; 350 C; 447 G; 523 T; 0 other:

Query Match      84.0%; Score 16.8; DB 22; Length 1954;
Best Local Similarity 90.0%; Pred. No. 1.5e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

{
  0Y      1 cctgtatgacctgcagaag 20
          11111111111111111111
  Db      200 ccagatagacctgcagaag 219

RESULT 11
AA114982
ID AA114982 standard; DNA; 1955 BP.
XX
XX AA114982;
XX
XX 12-OCT-2001 (first entry)
XX
XX Probe #4915 for gene expression analysis in human cervical cell sample.
DE
XX
XX Probe: human; microarray; gene expression; cervical epithelial cell;
KM cervical cancer; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200157278-A2.
PN
XX
XX 09-AUG-2001.
PD
XX
XX 30-JAN-2001; 2001WO-US00670.
PF
XX
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0642466.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX
XX Penn SG, Hanzel DK, Chen W, Rank DK;
P1
XX
XX WPI: 2001-488901/53.
DR
XX
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human cervical epithelial cells -
XX
XX Claim 25: SEQ ID No 4915; 487pp; English.
XX
XX The present invention relates to human single exon nucleic acid probes
CC (SENPs). The present sequence is one such probe. The SENPs are derived
CC from human HeLa cells. The SENPs can be used to produce a single exon
CC microarray, which can be used for measuring human gene expression in a
CC sample derived from human cervical epithelial cells. By measuring gene
CC expression, the probes are therefore useful in grading and/or staging
CC of diseases of the cervix, notably cervical cancer.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1955 BP; 714 A; 366 C; 526 G; 349 T; 0 other:

```

Query Match 84.0%; Score 16.8; DB 22; Length 1955;  
 Best Local Similarity 90.0%; Pred. No. 1.5e+02;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 cctgtatgacctgcagaag 20  
 || |||||  
 DB 1641 ccagtatgacctgcagaag 1660

## RESULT 12

AA136328  
 ID AA136328 standard; DNA: 1955 BP.

XX AA136328;

XX 17-OCT-2001 (first entry)

DE Probe #5014 used to measure gene expression in human placenta sample.

XX Probe; microarray: human; placenta; antenatal diagnosis;  
 XX genetic disorder; ss.

XX Homo sapiens.

XX WO200157272-A2.

XX 09-AUG-2001.

XX 30-JAN-2001: 2001WO-US00663.

XX 04-FEB-2000: 2000US-0180312.

XX 26-MAY-2000: 2000US-0207456.

XX 30-JUN-2000: 2000US-0608408.

XX 03-AUG-2000: 2000US-0632366.

XX 21-SEP-2000: 2000US-0234687.

XX 27-SEP-2000: 2000US-0236359.

XX 04-OCT-2000: 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SC, Hanzel DK, Chen W, Rank DR;

XX WPI: 2001-488897/53.

XX Human genome-derived single exon nucleic acid probes useful for

XX analyzing gene expression in human placenta -

XX Claim 25: SEQ ID No 5014; 654bp; English.

XX The present invention relates to single exon nucleic acid probes (SENP).

XX The present sequence is one such probe. The probes are useful for

XX producing a microarray for predicting, measuring and displaying gene

XX expression in samples derived from human placenta. The probes are useful

XX for antenatal diagnosis of human genetic disorders.

XX Sequence 1955 BP; 714 A; 366 C; 526 G; 349 T; 0 other;

XX Query Match 84.0%; Score 16.8; DB 22; Length 1955;

XX Best Local Similarity 90.0%; Pred. No. 1.5e+02;

XX Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 cctgtatgacctgcagaag 20

|| |||||  
 DB 1641 ccagtatgacctgcagaag 1660

## RESULT 13

AA104749  
 ID AA104749 standard; DNA: 1955 BP.

XX AA104749;

XX

DT 09-OCT-2001 (first entry)

XX Probe #4740 used to measure gene expression in human breast sample.

DE Probe; human; breast disease; breast cancer; development disorder; ss;

XX inflammatory disease; proliferative breast disease; non-carcinoma tumour.

XX Homo sapiens.

XX WO200157270-A2.

XX 09-AUG-2001.

XX 29-JAN-2001: 2001WO-US00661.

XX 04-FEB-2000: 2000US-0180312.

XX 26-MAY-2000: 2000US-0207456.

XX 30-JUN-2000: 2000US-0608408.

XX 03-AUG-2000: 2000US-0632366.

XX 21-SEP-2000: 2000US-0234687.

XX 27-SEP-2000: 2000US-0236359.

XX 04-OCT-2000: 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SC, Hanzel DK, Chen W, Rank DR;

XX WPI: 2001-476286/51.

XX Novel single exon nucleic acid probe used to measuring gene expression

XX in a human breast -

XX Claim 25: SEQ ID No 4740; 322bp; English.

XX The present invention relates to novel single exon nucleic acid probes.

XX The present sequence is one such probe. The probes are useful for

XX measuring human gene expression in a human breast sample, where the probe

XX hybridises at high stringency to a nucleic acid expressed in the human

XX breast. The probes are useful for predicting, diagnosing, grading,

XX staging, monitoring and prognosing diseases of the human breast.

XX particularly those diseases with polygenic aetiology. The diseases

XX of the breast, fibrocystic changes, proliferative breast disease and

XX non-carcinoma tumours.

XX Note: The sequence data for this patent did not form part of the printed

XX specification, but was obtained in electronic format directly from WIPO

XX at <http://wipo.int/patdb/patdbseqs.html>.

XX Sequence 1955 BP; 714 A; 366 C; 526 G; 349 T; 0 other;

XX Query Match 84.0%; Score 16.8; DB 22; Length 1955;

XX Best Local Similarity 90.0%; Pred. No. 1.5e+02;

XX Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 cctgtatgacctgcagaag 20

|| |||||  
 DB 1641 ccagtatgacctgcagaag 1660

## RESULT 14

AAH57390  
 ID AAH57390 standard; cDNA: 6016 BP.

XX AAH57390;

XX 10-SEP-2001 (first entry)

XX Human skeletal muscle cell specific cDNA sequence SEQ ID NO:230.

XX Human; tissue specific; diagnosis; brain; heart; skeletal muscle;

XX lung; liver; uterus; ovary; stomach; intestine; kidney; pancreas; ss;

XX metabolic disease; developmental disease; cytostatic; immunomodulatory;

XX

KW neuroprotective; gene therapy; cancer; immunopathology; neuropathology.  
 XX Homo sapiens.  
 OS  
 XX MO200132927-A2.  
 PN  
 XX 10-MAY-2001.  
 PD  
 XX 02-NOV-2000; 2000MO-US30396.  
 PF  
 XX 04-NOV-1999; 99US-0163508.  
 PR  
 XX (INCYTE GENOMICS INC.  
 PA  
 XX Sornasse T, Seilhamer JJ, Watson GA;  
 P1  
 XX WPI; 2001-291057/30.  
 DR  
 XX New cell and tissue specific polynucleotides useful for diagnosis,  
 PT prognosis or monitoring of treatments for disorders where the gene is  
 associated with a cancer, immunopathology or neuropathology -  
 PS  
 XX Claim 1; Page 158-159; 327pp; English.  
 CC AAH57161 to AAH57576 represent cell and tissue specific polynucleotide  
 CC sequences (I). (I) can have cytostatic, immunomodulatory and  
 CC neuroprotective activities, and can be used in gene therapy. (I) and  
 CC proteins (II) encoded by them are used in high throughput screening  
 CC assays to select DNA molecules, RNA molecules, peptide nucleic acids,  
 CC mimetics, peptides, proteins, agonists, antagonists, antibodies or  
 CC their fragments, immunoglobulins, inhibitors, drug compounds and  
 CC pharmaceutical agents. Expression of (I) in a sample indicates the  
 CC differentiation of embryonic stem cells into a tissue selected from  
 CC brain, heart, kidney, liver, lung, skeletal muscle or pancreatic  
 CC tissues. (I) and (II) are used to produce an expression profile that  
 CC defines a metabolic or developmental process, treatment, condition,  
 CC disease or disorder. The gene profile can be used for diagnosis,  
 CC prognosis or monitoring of treatments and for investigating a  
 CC predisposition to a disorder where the gene is associated with a  
 CC cancer, immunopathology or neuropathology.  
 CC  
 XX  
 SQ Sequence 6016 BP; 1953 A; 1297 C; 1590 G; 1176 T; 0 other;

Query Match 84.0%; Score 16.8; DB 22; Length 6016;  
 Best Local Similarity 90.0%; Pred. No. 1.6e+02;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Q 1 cctatagacttgcaggaag 20  
 ||| ||||| ||||| ||  
 DB 3100 cctgagatgacctgcagcagcag 3119

RESULT 15  
 AAZ24616  
 ID AAZ24616 standard; cDNA; 8948 BP.  
 XX  
 AC AAZ24616;  
 XX  
 UF 07-DEC-1999 (first entry)  
 DE Human lung tumor associated polynucleotide.  
 XX  
 KW Human; lung tumor; lung cancer; T cell stimulation; ss.  
 OS  
 XX Homo sapiens.  
 PN WO9947674-A2.  
 XX  
 PD 23-SEP-1999.  
 XX  
 PF 17-MAR-1999; 99WO-US05798.  
 XX

PR 18-MAR-1998; 98US-0040802.  
 PR 18-MAR-1998; 98US-0040984.  
 PR 27-JUL-1998; 98US-0123912.  
 PR 27-JUL-1998; 98US-0123933.  
 XX  
 PA (CORI-) CORIXA CORP.  
 XX  
 PI Reed SG, Wang T;  
 DR WPI; 1999-571839/48.  
 XX  
 XX New isolated lung tumor polynucleotides, used to develop products for  
 PT the treatment, prevention and monitoring the progression of lung cancer  
 PS  
 XX Claim 12; Page 111-114; 148pp; English.  
 CC  
 CC The invention provides isolated human lung tumor nucleic acids and  
 CC polypeptides. The polypeptides can be used for the treatment of lung  
 CC cancer. The polypeptides and polynucleotides can be used to stimulate T  
 CC cells or antigen presenting cells for use in the treatment of lung  
 CC cancer. The polypeptides and monoclonal antibodies specific for the  
 CC polypeptides can also be used to inhibit the development of lung cancer.  
 CC Agents which bind the polypeptides can be used for detecting lung cancer  
 CC and for monitoring the progression of lung cancer.  
 CC  
 XX  
 SQ Sequence 8948 BP; 2786 A; 2037 C; 2367 G; 1758 T; 0 other;

Query Match 84.0%; Score 16.8; DB 20; Length 8948;  
 Best Local Similarity 90.0%; Pred. No. 1.7e+02;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 cctatagacttgcaggaag 20  
 ||| ||||| ||||| ||  
 DB 4862 cctatagacttgcaggaag 4881

Search completed: March 4, 2002, 05:21:26  
 Job time: 17515 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 4, 2002, 05:11:33 : Search time 226.99 Seconds  
(without alignments)  
19,955 Million cell updates/sec

Title: US-08-956-991-6

Perfect score: 20

Sequence: 1 cctgtagacclycaggaag 20

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 351203 seqs, 113238999 residues

1 number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08  
Maximum Match 1008

Listing first 45 summaries

Database : Issued\_Patents.NA.\*

1: /cgn2\_6/prodata/2/ina/5A.COMB.seq.\*  
2: /cgn2\_6/prodata/2/ina/5B.COMB.seq.\*  
3: /cgn2\_6/prodata/2/ina/6A.COMB.seq.\*  
4: /cgn2\_6/prodata/2/ina/6B.COMB.seq.\*  
5: /cgn2\_6/prodata/2/ina/PCTUS.COMB.seq.\*  
6: /cgn2\_6/prodata/2/ina/backfiles1.seq.\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	16.8	84.0	1438	3	US-09-187-331-4
2	15.8	79.0	840	4	US-08-998-416-530
3	15.8	79.0	1296	1	US-08-090-523-1
4	15.8	79.0	1296	1	US-08-090-523-3
5	15.8	79.0	1296	1	US-08-398-627-1
6	15.8	79.0	1296	1	US-08-398-627-3
7	15.8	79.0	1296	1	US-08-406-858-1
8	15.8	79.0	1296	1	US-08-406-858-3
9	15.8	79.0	1296	5	PCT-US91-04036-1
10	15.8	79.0	1296	5	PCT-US91-04036-3
11	15.8	79.0	1296	5	PCT-US94-05275-1
12	15.8	79.0	1296	5	PCT-US94-05275-3
13	15.8	79.0	1323	1	US-07-735-065-3
14	15.8	79.0	1323	1	US-08-469-202-13
15	15.8	79.0	1323	2	US-08-484-414C-13
16	15.8	79.0	1323	2	US-08-229-515A-14
17	15.8	79.0	3955	1	US-08-645-865-14
18	15.4	77.0	3955	1	PCT-US93-00321-2
19	15.4	77.0	24	5	PCT-US93-00321-4
20	15.4	77.0	5102	1	US-08-494-168-1
21	15.2	76.0	22	1	US-08-184-422-15
22	15.2	76.0	22	3	US-08-589-771B-15
23	15.2	76.0	671	4	US-09-182-117-26
24	15.2	76.0	818	4	US-09-363-970-39
25	15.2	76.0	1003	4	US-09-363-970-38
26	15.2	76.0	1186	1	US-08-064-121-2
27	15.2	76.0	1186	1	US-08-478-015-2

28	15.2	76.0	1186	3	US-08-475-975-2	Sequence 2, Appl1
29	15.2	76.0	1186	3	US-09-084-889-2	Sequence 2, Appl1
30	15.2	76.0	1415	1	US-08-257-341-8	Sequence 8, Appl1
31	15.2	76.0	1515	4	US-09-080-625-6	Sequence 6, Appl1
32	15.2	76.0	1521	3	US-08-646-538-25	Sequence 25, Appl1
33	15.2	76.0	1521	3	US-08-646-538-31	Sequence 31, Appl1
34	15.2	76.0	1521	4	US-09-503-222-25	Sequence 25, Appl1
35	15.2	76.0	1521	4	US-09-503-222-31	Sequence 31, Appl1
36	15.2	76.0	1701	4	US-09-080-625-1	Sequence 1, Appl1
37	15.2	76.0	1853	1	US-08-404-732A-6	Sequence 6, Appl1
38	15.2	76.0	1872	3	US-08-422-108-2	Sequence 2, Appl1
39	15.2	76.0	2119	3	US-09-032-372-6	Sequence 6, Appl1
40	15.2	76.0	2120	1	US-08-404-732A-4	Sequence 4, Appl1
41	15.2	76.0	2150	3	US-09-263-023-1	Sequence 1, Appl1
42	15.2	76.0	2220	2	US-08-864-224-1	Sequence 8, Appl1
43	15.2	76.0	2224	1	US-08-404-732A-8	Sequence 8, Appl1
44	15.2	76.0	2245	4	US-09-442-055-3	Sequence 3, Appl1
45	15.2	76.0	2385	2	US-09-146-283-3	Sequence 3, Appl1

#### ALIGNMENTS

```
RESULT 1
US-09-187-331-4
: Sequence 4, Application US/09187331
: Patent No. 6043056
: GENERAL INFORMATION:
: APPLICANT: Yue, Henry
: APPLICANT: Corley, Neil C.
: APPLICANT: Guegler, Karl J.
: APPLICANT: Gorgone, Gina A.
: APPLICANT: Baughn, Mariah R.
: TITLE OF INVENTION: CELL SURFACE GLYCOPROTEINS
: FILE REFERENCE: PR-0631 US
: CURRENT APPLICATION NUMBER: US/09/187,331
: CURRENT FILING DATE: 1998-11-06
: NUMBER OF SEQ ID NOS: 6
: SOFTWARE: PERL Program
: SEQ ID NO 4
: LENGTH: 1438
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: 2705267
US-09-187-331-4

Query Match      84.0%  Score 16.8:  DB 3:  Length 1438:
Best Local Similarity 90.0%:  Pred. No. 17:
Matches 18:  Conservative 0:  Mismatches 2:  Indels 0:  Gaps 0:

QY      1 cctgtagacclycaggaag 20
Db      863 cctgtagacclycaggaag 882

RESULT 2
US-08-998-416-530/c
: Sequence 530, Application US/08998416
: Patent No. 6239264
: GENERAL INFORMATION:
: APPLICANT: Philippson, Peter
: APPLICANT: Pohlmann, Rainer
: APPLICANT: Steiner, Sabine
: APPLICANT: Mohr, Christine
: APPLICANT: Wendland, Jurgen
: APPLICANT: Knechtle, Philipp
: APPLICANT: Reischung, Corinne
: TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSYPTII
: TITLE OF INVENTION: AND USES THEREOF
: NUMBER OF SEQUENCES: 1152
: CORRESPONDENCE ADDRESSES:
```

ADDRESSEE: No. 6239264rtis Corporation  
STREET: 3054 Cornwallis Road  
CITY: Research Triangle Park  
STATE: No. 6239264th Carolina  
COUNTRY: USA  
ZIP: 27709  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/998,416  
FILING DATE: 24-DEC-1997  
CLASSIFICATION: 435  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: CH 0016/97  
FILING DATE: 31-DEC-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Meigs, J. Timothy  
REGISTRATION NUMBER: 38,241  
REFERENCE/DOCKET NUMBER: PF/5-30306/A/COC1976  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-541-8587  
TELEFAX: 919-541-8589  
INFORMATION FOR SEQ ID NO: 530:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 840 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE:  
ORGANISM: PAC1370UP  
US-08-998-416-530

Query Match 79.0%; Score 15.8; DB 4; Length 840;  
Best Local Similarity 89.5%; Pred. No. 47;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ctgataagacctgcaggag 20  
|||||  
DB 607 CTGTATGACCTGCTGCAG 589

RESULT 3  
US-090-523-1  
Invention No. 5498830  
GENERAL INFORMATION:  
APPLICANT: Barry, Gerard F.  
APPLICANT: Kishore, Ganesh M.  
APPLICANT: Stark, David M.  
TITLE OF INVENTION: Enhanced starch biosynthesis  
NUMBER OF SEQUENCES: 51  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Grace L. Bonner, Monsanto Co. BBAF  
STREET: 700 Chesterfield Parkway No. 5498830th  
CITY: St. Louis  
STATE: Missouri  
COUNTRY: USA  
ZIP: 63198  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/090,523  
FILING DATE: 19930712  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/709663  
FILING DATE: 07-JUN-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/539763  
FILING DATE: 18-JUN-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Bonner, Grace L.  
REGISTRATION NUMBER: 32,963  
REFERENCE/DOCKET NUMBER: 38-21(10559)A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (314) 537-7286  
TELEFAX: (314) 537-6047  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1296 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1293  
US-08-090-523-1

Query Match 79.0%; Score 15.8; DB 1; Length 1296;  
Best Local Similarity 89.5%; Pred. No. 49;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ctgataagacctgcaggag 20  
|||||  
DB 667 CTGTATGACCTGCTGCAG 685

RESULT 4  
US-08-090-523-1  
Invention No. 5498830  
GENERAL INFORMATION:  
APPLICANT: Barry, Gerard F.  
APPLICANT: Kishore, Ganesh M.  
APPLICANT: Stark, David M.  
TITLE OF INVENTION: Enhanced starch biosynthesis  
NUMBER OF SEQUENCES: 51  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Grace L. Bonner, Monsanto Co. BBAF  
STREET: 700 Chesterfield Parkway No. 5498830th  
CITY: St. Louis  
STATE: Missouri  
COUNTRY: USA  
ZIP: 63198  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/090,523  
FILING DATE: 19930712  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/709663  
FILING DATE: 07-JUN-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/539763  
FILING DATE: 18-JUN-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Bonner, Grace L.  
REGISTRATION NUMBER: 32,963  
REFERENCE/DOCKET NUMBER: 38-21(10559)A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (314) 537-7286  
TELEFAX: (314) 537-6047

INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1296 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1293  
US-08-090-523-3

Query Match 79.0%; Score 15.8; DB 1; Length 1296;  
Best Local Similarity 89.5%; Pred. No. 49;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ctgatagacctgcaggaag 20  
||||| ||| ||| |||  
D 667 CTGTATGAACCTGCTGGAG 685

RESULT 5  
US-08-398-627-1  
Sequence 1, Application US/08398627  
Patent No. 5608149  
GENERAL INFORMATION:  
APPLICANT: Barry, Gerard F.  
APPLICANT: Kishore, Ganesh M.  
APPLICANT: Stark, David M.  
TITLE OF INVENTION: Enhanced Starch Biosyntheses  
NUMBER OF SEQUENCES: 51  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Grace L. Bonner, Monsanto Co. B44F  
STREET: 700 Chesterfield Parkway No. 5608149th  
CITY: St. Louis  
STATE: Missouri  
COUNTRY: USA  
ZIP: 63198  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/398.627  
FILING DATE: 03-MAR-1995  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/090.523  
FILING DATE: 12-JUL-1993  
APPLICATION NUMBER: US 07/709663  
FILING DATE: 07-JUN-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/539763  
FILING DATE: 18-JUN-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Bonner, Grace L.  
REGISTRATION NUMBER: 32,963  
REFERENCE/DOCKET NUMBER: 38-21(10559)A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (314) 537-7286  
TELEFAX: (314) 537-6047  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1296 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1293

US-08-398-627-1

Query Match 79.0%; Score 15.8; DB 1; Length 1296;  
Best Local Similarity 89.5%; Pred. No. 49;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ctgatagacctgcaggaag 20  
||||| ||| ||| |||  
D 667 CTGTATGAACCTGCTGGAG 685

RESULT 6  
US-08-398-627-3  
Sequence 3, Application US/08398627  
Patent No. 5608149  
GENERAL INFORMATION:  
APPLICANT: Barry, Gerard F.  
APPLICANT: Kishore, Ganesh M.  
APPLICANT: Stark, David M.  
TITLE OF INVENTION: Enhanced Starch Biosyntheses  
NUMBER OF SEQUENCES: 51  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Grace L. Bonner, Monsanto Co. B44F  
STREET: 700 Chesterfield Parkway No. 5608149th  
CITY: St. Louis  
STATE: Missouri  
COUNTRY: USA  
ZIP: 63198  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/398.627  
FILING DATE: 03-MAR-1995  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/090.523  
FILING DATE: 12-JUL-1993  
APPLICATION NUMBER: US 07/709663  
FILING DATE: 07-JUN-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/539763  
FILING DATE: 18-JUN-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Bonner, Grace L.  
REGISTRATION NUMBER: 32,963  
REFERENCE/DOCKET NUMBER: 38-21(10559)A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (314) 537-7286  
TELEFAX: (314) 537-6047  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1296 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1293  
US-08-398-627-3

Query Match 79.0%; Score 15.8; DB 1; Length 1296;  
Best Local Similarity 89.5%; Pred. No. 49;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ctgatagacctgcaggaag 20  
||||| ||| ||| |||  
D 667 CTGTATGAACCTGCTGGAG 685

RESULT 7  
US-08-406-858-1  
Sequence 1, Application US/08406858  
Patent No. 5648249  
GENERAL INFORMATION:  
APPLICANT: Barry, Gerard F.  
APPLICANT: Kishore, Ganesh M.  
APPLICANT: Stark, David M.  
APPLICANT: Zaleski, James C.  
TITLE OF INVENTION: Method of improving the quality of  
TITLE OF INVENTION: Stored Potatoes  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Grace L. Bonner, Monsanto Company, B84F  
STREET: 700 Chesterfield Parkway No. 5648249th  
CITY: St. Louis  
STATE: Missouri  
COUNTRY: USA  
ZIP: 63198  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/406,858  
FILING DATE:  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/05275  
FILING DATE: 18-MAY-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/070,155  
FILING DATE: 28-MAY-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Bonner, Grace L.  
REGISTRATION NUMBER: 32,963  
REFERENCE/DOCKET NUMBER: 38-21(10654)A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (314)537-7286  
TELEFAX: (314)537-6047  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1296 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1296  
US-08-406-858-1

Query Match 79.0%; Score 15.8; DB 1; Length 196;  
Best Local Similarity 89.5%; Pred. No. 49;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ctglatgacctgcagaaq 20  
||||||| ||||| |||||  
db 667 CTGTATGAACTGCTGGAAG 685

RESULT 8  
US-08-406-858-3  
Sequence 3, Application US/08406858  
GENERAL INFORMATION:  
APPLICANT: Barry, Gerard F.  
APPLICANT: Kishore, Ganesh M.  
APPLICANT: Stark, David M.

APPLICANT: Zaleski, James C.  
TITLE OF INVENTION: Method of improving the quality of  
TITLE OF INVENTION: Stored Potatoes  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Grace L. Bonner, Monsanto Company, B84F  
STREET: 700 Chesterfield Parkway No. 5648249th  
CITY: St. Louis  
STATE: Missouri  
COUNTRY: USA  
ZIP: 63198  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/406,858  
FILING DATE:  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/05275  
FILING DATE: 18-MAY-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/070,155  
FILING DATE: 28-MAY-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Bonner, Grace L.  
REGISTRATION NUMBER: 32,963  
REFERENCE/DOCKET NUMBER: 38-21(10654)A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (314)537-7286  
TELEFAX: (314)537-6047  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1296 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1296  
US-08-406-858-4

Query Match 79.0%; Score 15.8; DB 1; Length 1296;  
Best Local Similarity 89.5%; Pred. No. 49;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ctglatgacctgcagaaq 20  
||||||| ||||| |||||  
db 667 CTGTATGAACTGCTGGAAG 685

RESULT 9  
PCT-US91-004036-1  
Sequence 1, Application PCT/US9104036  
GENERAL INFORMATION:  
APPLICANT: Kishore, Ganesh M.  
TITLE OF INVENTION: Increased starch content in plants  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Monsanto Co.  
STREET: 700 Chesterfield Village Parkway  
CITY: St. Louis  
STATE: Missouri  
COUNTRY: USA  
ZIP: 63198  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA: PCT/US91/04036  
APPLICATION NUMBER: PCT/US91/04036  
FILING DATE: 19910607  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: McBride, Thomas P.  
REGISTRATION NUMBER: 32706  
REFERENCE/DOCKET NUMBER: 38-21(10530)A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (314) 537-7357  
TELEFAX: (314) 537-6047  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1296 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1293  
PCT-US91-04036-1

Query Match 79.0%; Score 15.8; Db 5; Length 1296,  
Best Local Similarity 89.5%; Pred. No. 49;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 cgtatgaccctgcgaag 20  
||||| 1111 1111  
db 667 CTGTATGACCTGCTGGAAG 685

RESULT 10  
PCT-US91-04036-3  
Sequence 3, Application PC/TUS9104036  
GENERAL INFORMATION:  
APPLICANT: Kishore, Ganesh M.  
TITLE OF INVENTION: Increased Starch Content in Plants  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Monsanto Co.  
STREET: 700 Chesterfield Village Parkway  
CITY: St. Louis  
STATE: Missouri  
COUNTRY: USA  
ZIP: 63198  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US91/04036  
FILING DATE: 19910607  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: McBride, Thomas P.  
REGISTRATION NUMBER: 32706  
REFERENCE/DOCKET NUMBER: 38-21(10530)A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (314) 537-7357  
TELEFAX: (314) 537-6047  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1296 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS

LOCATION: 1..1293  
PCT-US91-04036-3

Query Match 79.0%; Score 15.8; DB 5; Length 1296,  
Best Local Similarity 89.5%; Pred. No. 49;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 cgtatgaccctgcgaag 20  
||||| 1111 1111  
db 667 CTGTATGACCTGCTGGAAG 685

RESULT 11  
PCT-US94-05275-1  
Sequence 1, Application PC/TUS9405275  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: Method of Improving the Quality of Stored  
NUMBER OF SEQUENCES: 26  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/05275  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
PCT-US94-05275-3

Query Match 79.0%; Score 15.8; DB 5; Length 1296,  
Best Local Similarity 89.5%; Pred. No. 49;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 cgtatgaccctgcgaag 20  
||||| 1111 1111  
db 667 CTGTATGACCTGCTGGAAG 685

RESULT 12  
PCT-US94-05275-3  
Sequence 3, Application PC/TUS9405275  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: Method of Improving the Quality of Stored  
NUMBER OF SEQUENCES: 26  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/05275  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:



APPLICATION NUMBER: US 08/070155  
FILING DATE: 28-MAY-1993  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1296 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1296  
PCT-US94-05275-3

Query Match 79.0% Score 15.8; DB 5; Length 1296;  
Best Local Similarity 89.5% Pred. No. 49;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

2 ctglatgacctgcaggaag 20  
||||| ||| ||| |||  
db . 667 CTGTATGACTGCTGGAAG 685

RESULT 13  
US-07-735-065-3  
Sequence 3, Application US/07735065  
Patent No. 5349123  
GENERAL INFORMATION:  
APPLICANT: Christine K. Shewmaker  
APPLICANT: David M. Stalker  
TITLE OF INVENTION: Glycogen Biosynthetic Enzymes  
TITLE OF INVENTION: In Plants  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Calgene, Inc.  
STREET: 1920 Fifth Street  
CITY: Davis  
STATE: CA  
COUNTRY: USA  
ZIP: 95616  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.0 MB  
COMPUTER: Apple Macintosh  
OPERATING SYSTEM: Macintosh 6.0.7  
SOFTWARE: Microsoft Word 4.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/735,065  
FILING DATE: 19910724  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/632,383  
FILING DATE: 21-DEC-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: unassigned  
FILING DATE: 16-JUL-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Elizabeth Lassen  
REGISTRATION NUMBER: 31,845  
NAME: Donna E. Scherer  
REGISTRATION NUMBER: 34,719  
REFERENCE/DOCKET NUMBER: CGNE 77-2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (916) 753-6313  
TELEFAX: (916) 753-1510  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1323 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: genomic DNA  
US-07-735-065-3

Query Match 79.0% Score 15.8; DB 1; Length 1323;  
Best Local Similarity 89.5% Pred. No. 49;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
2 ctglatgacctgcaggaag 20  
||||| ||| ||| |||  
db 682 CTGTATGACTGCTGGAAG 700

RESULT 14  
US-08-469-202-13  
Sequence 13, Application US/08469202  
Patent No. 5750875  
GENERAL INFORMATION:  
APPLICANT: STALKER, DAVID  
TITLE OF INVENTION: GLYCOGEN BIOSYNTHETIC ENZYMES IN PLANTS  
NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CALGENE, INC.  
STREET: 1920 FIFTH STREET  
CITY: DAVIS  
STATE: CA  
COUNTRY: USA  
ZIP: 95616  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB  
COMPUTER: Apple Macintosh  
OPERATING SYSTEM: Macintosh 7.1  
SOFTWARE: Microsoft Word 5.1 (a)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/469,202  
FILING DATE: 6-JUNE-95  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/016,881  
FILING DATE: 11-FEB-1993  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Elizabeth Lassen  
REGISTRATION NUMBER: 31,845  
NAME: Donna E. Scherer  
REGISTRATION NUMBER: 34,719  
NAME: Carl J. Schwedler  
REGISTRATION NUMBER: 36,924  
REFERENCE/DOCKET NUMBER: CGNE 93-2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 916-753-6313  
TELEFAX: 916-753-1510  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1323 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
US-08-469-202-13

Query Match 79.0% Score 15.8; DB 1; Length 1323;  
Best Local Similarity 89.5% Pred. No. 49;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

2 ctglatgacctgcaggaag 20  
||||| ||| ||| |||  
db 682 CTGTATGACTGCTGGAAG 700

RESULT 15  
US-08-484-434C-13  
Sequence 13, Application US/08484434C

Patent No. 5969214  
GENERAL INFORMATION:  
APPLICANT: STALKER, DAVID  
TITLE OF INVENTION: GLYCOSYL BIOSYNTHETIC ENZYMES IN PLANTS  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Calgene LLC  
STREET: 1920 Fifth Street  
CITY: Davis  
STATE: CA  
COUNTRY: USA  
ZIP: 95616  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC  
OPERATING SYSTEM: Windows NT 4.0; MS Word for Windows 7.0a  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,434C  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/016,881  
FILING DATE: 11-FEB-1993  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Schwedler, Carl J.  
REGISTRATION NUMBER: 36,924  
REFERENCE/DOCKET NUMBER: 15593/01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 530-792-2265  
TELEFAX: 530-792-2463  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1323 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 16..1308  
US-08-484-434C-13

Query Match 79.0%; Score 15.8; DB 2; Length 1323;  
Best Local Similarity 89.5%; Pred. No. 49;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

U 2 cgtatgacctcaggaag 20  
|||||  
DB 682 cgtatgacctcaggaag 700

Search completed: March 4, 2002, 05:11:34  
Job time: 16948 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 4, 2002, 05:21:26 : Search time 578.64 Seconds  
(without alignments)  
1247.525 Million cell updates/sec

Title: US-08-956-991-7

Perfect score: 842  
Sequence: 1 ccggggccggcgcgcgag.....cgcgtgagctgccttcca 842

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 930621 seqs, 428662619 residues  
number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

1: N\_Geneseq\_1101.\*  
2: /SID52/gcgdata/geneseq/geneseqn/NA1980.DAT.\*  
3: /SID52/gcgdata/geneseq/geneseqn/NA1981.DAT.\*  
4: /SID52/gcgdata/geneseq/geneseqn/NA1982.DAT.\*  
5: /SID52/gcgdata/geneseq/geneseqn/NA1983.DAT.\*  
6: /SID52/gcgdata/geneseq/geneseqn/NA1984.DAT.\*  
7: /SID52/gcgdata/geneseq/geneseqn/NA1985.DAT.\*  
8: /SID52/gcgdata/geneseq/geneseqn/NA1986.DAT.\*  
9: /SID52/gcgdata/geneseq/geneseqn/NA1987.DAT.\*  
10: /SID52/gcgdata/geneseq/geneseqn/NA1988.DAT.\*  
11: /SID52/gcgdata/geneseq/geneseqn/NA1989.DAT.\*  
12: /SID52/gcgdata/geneseq/geneseqn/NA1990.DAT.\*  
13: /SID52/gcgdata/geneseq/geneseqn/NA1991.DAT.\*  
14: /SID52/gcgdata/geneseq/geneseqn/NA1992.DAT.\*  
15: /SID52/gcgdata/geneseq/geneseqn/NA1993.DAT.\*  
16: /SID52/gcgdata/geneseq/geneseqn/NA1994.DAT.\*  
17: /SID52/gcgdata/geneseq/geneseqn/NA1995.DAT.\*  
18: /SID52/gcgdata/geneseq/geneseqn/NA1996.DAT.\*  
19: /SID52/gcgdata/geneseq/geneseqn/NA1997.DAT.\*  
20: /SID52/gcgdata/geneseq/geneseqn/NA1998.DAT.\*  
21: /SID52/gcgdata/geneseq/geneseqn/NA1999.DAT.\*  
22: /SID52/gcgdata/geneseq/geneseqn/NA2000.DAT.\*  
23: /SID52/gcgdata/geneseq/geneseqn/NA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	842	100.0	842	19	AAV31985
2	670.2	79.6	6413	19	AAV31988
3	670.2	79.6	6604	19	AAV31981
4	307.6	36.5	1493	19	AAV27207
5	76.2	9.0	5824	22	AA158196
6	39.2	4.7	4257	19	AAV68520
7	39.2	4.7	4257	19	AAV10362
8	39.2	4.7	7686	21	AAZ29335
9	39.2	4.7	12001	16	AAQ76213
10	38.4	4.6	3588	8	AAW70317
11	38	4.5	2900	13	AAQ20372

12	37.4	4.4	7191	21	AA97538
13	36.6	4.3	698	14	AA38759
14	36.6	4.3	3694	21	AAQ36930
15	36.4	4.3	1524	18	AA63571
16	36.2	4.3	731	22	AA644924
17	35.8	4.3	2641	22	AA161175
18	35.8	4.3	37856	21	AA11992
19	35.6	4.2	2886	20	AAZ09491
20	35.4	4.2	405	20	AA89576
21	35.4	4.2	4743	21	AA91408
22	35.2	4.2	6728	21	AA35248
23	35.2	4.2	109519	22	AA508693
24	35.2	4.2	725	22	AAH06483
25	35	4.2	2312	22	AAH16152
26	34.8	4.1	3608	10	AA91777
27	34.8	4.1	3802	16	AA87919
28	34.8	4.1	3802	18	AA76812
29	34.8	4.1	3802	18	AA70227
30	34.8	4.1	3802	21	AA258942
31	34.8	4.1	50341	19	AAV22674
32	34.8	4.1	50341	19	AA239519
33	34.8	4.1	52297	16	AA751411
34	34.8	4.1	52298	14	AAQ47357
35	34.4	4.1	1910	19	AAV48148
36	34.4	4.1	10097	22	AA697856
37	34.4	4.1	121162	21	AA66548
38	34.2	4.1	392	20	AA56151
39	34.2	4.1	457	20	AA56144
40	34.2	4.1	871	22	AAH06555
41	34.2	4.1	1170	13	AAQ20217
42	34.2	4.1	1548	19	AAV35840
43	34.2	4.1	1550	13	AAQ31550
44	34.2	4.1	1550	19	AAV41719
45	34.2	4.1	1550	20	AA78495

#### ALIGNMENTS

RESULT 1  
AAV31985  
ID AAV31985 standard: cDNA: 842 BP.  
XX  
AC AAV31985:  
XX  
DT 28-SEP-1998 (first entry)  
XX  
DE Mouse Down syndrome-cell adhesion molecule DS-CAM 5' cDNA.  
XX  
KW DS-CAM; Down syndrome-cell adhesion molecule; neural cell;  
KW signal transduction; trisomy 21; mental retardation;  
KW holoprosencephaly; corpus callosum agenesis;  
KW schizophrenia; diagnosis; assay; mouse; ds; ss.  
XX  
OS Mus sp.  
XX  
PN W09817795-A1.  
XX  
PD 30-APR-1998.  
XX  
PF 23-OCT-1997: 97W0-US19547.  
XX  
PR 25-OCT-1996: 96US-0029322.  
XX  
PA (CEDA-) CEDARS SINAI MEDICAL CENT.  
XX  
PI korenberg JR:  
XX  
DR WPI: 1998-271791/24.  
XX  
KT New isolated Down's Syndrome-cell adhesion molecule - used to  
PT develop products for detection, diagnosis and therapy of  
PT developmental and neurological abnormalities









[illegible][illegible]





















```

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-24-257g14"
/clone_lib="RPCI-24"
/sex="Male"
/cell_type="Spleen/Brain"
/note="Vector: pTRABAC1; Site_1: BamHI; Site_2: BamHI;
RPCI-24 Mouse BAC Library produced by Pieter de Jong. The
library was cloned in the pTRABAC1 cloning vector at the
BamHI sites using MboI partially digested male C57BL/6J
DNA."
BASE COUNT      222 a      141 c      154 g      193 t
ORIGIN

```

Query Match	17.0%:	Score 142.8:	DB 13:	Length 710:
Best Local Similarity	92.6%:	Pred. No. 1.8e-25:		
Matches 150:	Conservative	0:	Mismatches 12:	Indels 0: Gaps 0
Dn	598	t t c a c t g t c c a g a c t a g a t t t c a t c a a c t c a c g a g a g c c t t g a t a t a a y	657	
Oy	521	T T A T A T C T G T T C A G G A T T C T A G A T T T C T A T C A T C C A C G G A C C T T G A T A T T A A A	462	
Dn	658	a l t t c a a c a a c a a g a t g g a c t i g l a a c a c t a c c g t c a a l e g c g c g a c a g a t t c p g g	717	
Oy	461	A T G T T C A C A A C C A A G A T G G C T G T A C A A C A C C G C T G C A T C A C G G G A C A C A T A C A G	402	
Dn	718	g g g a g a c g a g a c a g a c a a c t g t g c g a g a c t t g t g c a g	759	
Oy	401	G G G A C A C G A C A C A A A G C A A C A C A G C G C A G A C T T T C T G T C A G	360	
RESULT	6			
HSMC18B12				
LOCUS	HSMC18B12	147 bp	DNA	GSS
DEFINITION	H.sapiens DNA for trapped exon (ID HMC18B12), genomic survey sequence.			
ACCESSION	X88325			
VERSION	X88325.1	GI:1437728		
KEYWORDS	GSS:			
SOURCE	human.			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
REFERENCE	1 (bases 1 to 147)			
AUTHORS	Chen,H.M., Rossier,C., Christ,R. and Antonarakis,S.E.			
JOURNAL	Cloning of trapped exons from human chromosome 21			
TITLE	Unpublished			
RENCE	2 (bases 1 to 147)			
AUTHORS	Antonarakis,S.E.			
TITLE	Direct Submission			
JOURNAL	Submitted (17-MAR-1995) Stylianou E, Antonarakis S, Division of Medical Genetics, University and Cantonal Hospital of Geneva, CMU 1 rue Michel-Servet, 1211 Geneva, SWITZERLAND			
REFERENCE	3 (bases 1 to 147)			
AUTHORS	Chen,H., Christ,R., Rossier,C., Morris,M.A., Lalioti,M.D. and Antonarakis,S.E.			
TITLE	Cloning of 559 potential exons of genes of human chromosome 21 by exon trapping			
JOURNAL	Genome Res. 6 (8), 747-760 (1996)			
LINE	97011340			
FEATURES	Location/Qualifiers			
SOURCE	1..147			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
	/chromosome="21"			
	1..147			
	/note="trapped exon"			
BASE COUNT	33 a	33 c	39 g	38 t
ORIGIN				4 others

[illegible]

	RESULT	7	GSS	15-MAY-2000
Locus	CNS03BK1/C	535 bp	DNA	
Definition	Tetraodon nigroviridis genome survey sequence pUC-ori end of clone 01ZC20 of library G from Tetraodon nigroviridis, genomic survey sequence.			
Accession	AL236602	GI:7895737		
Version	AL236602.1			
Keywords	GSS; genome survey sequence;			
Source	Tetraodon nigroviridis.			
Organism	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorphia; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.			
Reference	1 (bases 1 to 535) Koeist,-crol-lins,H., Jalllon,O., Dasilva,C., Fizesmes,C., Fisher,C., Bonneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and Weissenbach,J. Characterization and repeat analysis of the compact genome of the freshwater putterfish Tetraodon nigroviridis			
Title	Unpublished			
Journal	2 (bases 1 to 535)			
Reference	Bernot,A., Fizesmes,C., Winkler,P., Brotlier,P., Quetier,F., Saurin,W. and Weissenbach,J. Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence			
Authors	Unpublished			
Journal	3 (bases 1 to 535)			
Reference	Genoscope.			
Authors	Submitted (12-Apr-2000) to the EMBL/Genbank/DDBJ databases			
Title	This sequence is a single-read and was generated as part of a large scale clone-and-sequence project of the Tetraodon nigroviridis genome. For more information, please take a look at <a href="http://www.genoscope.cns.fr/Tetraodon.location/Qualifiers">http://www.genoscope.cns.fr/Tetraodon. location/Qualifiers</a>			
Comment	1..535 /organism="tetraodon nigroviridis" /db_xref="taxon:99883" /clone="01ZC20" /clone_1fb="G" /note="Genoscope sequence ID : C0BG012BB10SP1-end ; pUC-Ori"			
Features				
Source				
Base Count	140 a	107 c	165 g	102 t      21 others
Origin				
Query Match	15.0%:	Score 126.6:	DB 13:	Length 535;
Best Local Similarity	73.0%:	Pred. No. 1.9e+21,		
Matches	222:	Conservative	5;	Mismatches 70; Indels 7; Gaps 5.
OY	148	ttttaagtagagaagcccaatccc--aggctcatcttgatcatagcacgtcgctgaagaagt	205	
DB	430	TTTTCTGTGAACATTCATTCCAGGGTTATATTTCATCATCAATGCATCTGCCAAGAAGCT	371	

[illegible][illegible]

[illegible]

Query Match	10.1%;	Score 85.4;	DB 13;	Length 1068;
Best Local Similarity	66.7%;	Pred. No. 5,4e-11;		
Matches 122;	Conservative	0;	Mismatches 61;	Indels 0; Gaps 0;
Oy	606	gtctcagatcctgatttcctcatcacatccacggaggaccttataataagaatglttcag	665	
Dd	642	gTTTCACTGTCGGCGGTTTCTCATCATCAGACAGCCGAGGCCCTCTACACTCTGTGATGTGCAG	701	
Oy	666	aacgaagatlygr.tylactaaciacgcgtgatcgcggcgacagatcgcggggagacg	725	
Dd	702	ATGAGGACGCGGCTGTACACTACAGATGCATGACCGCGCATCGTACACAGAGAAGACC	761	
Oy	726	agacagagcaacatcgcgcgcgaatcgttgctcgaagaccgaactcaagccatccatcc	785	
Dd	762	CGGCAGAGCAACACGCGCCGCTTCATCTCTCACGTAGAGAGAGACAGACAGCACAGC	821	
Oy	786	tqg 788		
Dd	822	TGG 824		
RESULT 11				
CNS02130	862 bp	DNA	GSS	13-MAY-2000
LOCUS	Tetraodon nigroviridis genome survey sequence T7 end of clone			
DEFINITION	140P02 of library G from Tetraodon nigroviridis, genomic survey			
ACCESSION	AL198405.1	GI:7836556		
VERSION	ALI98405.1	CI:7836556		
KEYWORDS	GSS: genome survey sequence.			
SOURCE	Tetraodon nigroviridis.			
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorphi; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.			
REFERENCE	1 (bases 1 to 862)			
AUTHORS	Roeft-Crolijns,H., Jallion,O., Dasilva,C., Fizames,C., Fisher,C., Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and Weissenbach,J.			
TITLE	Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis			
JOURNAL	Unpublished			
REFERENCE	2 (bases 1 to 862)			
AUTHORS	Roeft-Crolijns,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fizames,C., Winkler,P., Brottier,P., Quetier,F., Saurin,W. and Weissenbach,J.			
TITLE	Human gene number estimation provided by genome wide analysis using Tetraodon nigroviridis DNA sequence			
JOURNAL	Unpublished			
REFERENCE	3 (bases 1 to 862)			
AUTHORS	Genoscope.			
TITLE	Direct Submission			
JOURNAL	Submitted (12-Apr-2000) to the EMBL/GenBank/DBJ databases			
COMMENT	This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at <a href="http://www.genoscope.crs.fr/tetraodon.location/Qualities">http://www.genoscope.crs.fr/tetraodon. location/Qualities</a>			
FEATURES				
source	1..862			
	/organism="Tetraodon nigroviridis"			
	/db_xref="taxon:99883"			
	/clone="140P02"			
	/clone_lib="G"			
	/note="Genoscope sequence ID : CGAG140DC01UP1-end : T7"			
BASE COUNT	269 a 154 c 204 g 212 t 23 others			
ORIGIN				
Query Match	9.4%;	Score 78.8;	DB 13;	Length 862;
Best Local Similarity	69.8%;	Pred No. 2.3e-09;		

Matches	104;	Conservative	1;	Mismatches	49;	Indels	0;	Gaps	0;
OY	465	gltttacggagacctatcacgtccglttgtagagctcgaataacatqayagyaatgltc		524					
Db	785	gTTTTCAAGGAGCCCTACACGCGTGCGGGGCGCGCGATCGGCGCTGAGGCAACCTT		726					
OY	525	gcgggtgttcacagtcatatccccctccctcgtggtgagggaglaqgtctgtgtaactcatg		584					
Db	725	GCTGTGTTCAAGTGCCATCCCGGCTGCTGTGCACAGGAATATATCAGCGTGTCTCTCG		666					
OY	585	gagaaagacacaggttcaactgtgtcagg		613					
Db	665	AAGAGACACACCTMTCTCATCTGCTCCAGG		637					

RESULT	12				
CNS02C1Y				GSS	12-MAY-2000
Locus	611 bp	DNA			
DEFINITION	Tetradon nigroviridis genome survey sequence pUC-ori end of clone				
	25N21 of library G from Tetradon nigroviridis, genomic survey sequence.				

ACCESSION	AL191311
VERSION	AL191311.1
KEYWORDS	GI:7829415
SOURCE	GSS: genome survey sequence, Tetradon nigriviridis, Tetradon nigriviridis, Tetradon nigriviridis

**ORGANISM**    *Nysius*: Hemiptera: Pentatomidae; Eukaryot.: Metazoa; Chordata; Craniata.; Vertebrata; Euteleostomi; Aclinopterygii; Neopterygii; Teleostei.; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes; Acanthomorpha; Tetraodon.; Tetraodontidae.

REFERENCE  
1 (bases 1 to 611)  
Roest-Crolius, H., Jallou, O., Dasilva, C., Fizames, C., Fisher, C.,  
Bonneau, L., Billault, A., Quelier, F., Saurin, W., Bernot, A. and  
Weissenbach, J. 1993. A detailed and repeat analysis of the compact genome of the

TITLE	REFERENCE
Unpublished	Reest-Crolius, H., Jallou, O., Basileva, C., Bouneau, L., Fisher, C., and Ouellet, F.
2 (basses 1 to 611)	
freshwater pufferfish Tetraodon nigroviridis	
tharaterization and reproduction	
JOURNAL	
AIRMOHS	

bernot, A., Fitzames, C., Winkler, P., Broctier, P., Voelckers, J., Saurin, W. and Weissbach, J.  
Human gene number estimate provided by genome wide analysis using  
Tetradon nigroviridis DNA sequence

JOURNAL, unpublished  
REFERENCE 3 (pages 1 to 611)  
AUTHORS Genoscope.  
TITLE Direct Submission  
Numbered (12-PPR-2000) to the EMBL/Genbank/DBI databases

This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetradon nitroviridis genome. For more information, please take a look at <http://www.genoscope.cns.fr/Tetradon>.

```

FEATURES
  source
    location/Qualifiers
      1..611
        /organism="Tetradodon nigroviridis"
        /db_xref="taxon:99883"
        /clonea="255N21"

```

BASE COLUMN	109 a	161 c	152 d	167 t	22 others
/clone_220266					
/clone_110="g"					
/note="Genoscope sequence ID : C0A6255C.G11SP1-end :					
PUC-Or1"					

Origin	Query Match
8.2%	Score 69; DB 13; Length 611;
5.7e-07;	Prod No. 5.7e-07;

[illegible][illegible]

Db 521 GTATTTCACAGCMCCCGGGGCGCTCGATGCTCTGTCGCCCGGGGAGCGAGCCCGCC 580

Oy 267 a..tctcagatggtacctagcaacgagcagag 297

Db 581 TCGCTCGCGTGGCTTCCTGGCCACCGCGCAGG 611

RESULT	13
CNS04L51	
LOCUS	CNS04L51 656 bp DNA GSS 21-MAY-2000
DEFINITION	Tetradon nigroviridis genome survey sequence T7 end of clone Tetradon nigroviridis genome survey

ACCESSION	AI295695
VERSION	AI295695.1
KEYWORDS	GI:8044275
SOURCE	GSS: genome survey sequence, Tetradodon nigroviridis.

ORGANISM  
Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:  
Actinopterygii: Neopterygii: Teleostei: Euteleostei: Neoteleostei  
Acanthomorphi: Acanthopterygii: Percomorpha: Tetraodontiformes;  
Tetraodontidae; Tetraodon

REFERENCE  
AUTHORS

1 (bases 1 to 656)  
Roest-Collins, H., Jallion, O., Davila, C., Fizames, C., Fisher, C.,  
Bonneau, L., Billault, A., Queller, F., Saurin, W., Bernot, A. and  
Weissenbach, J. 1994. A physical analysis of the compact genome of the

JOURNAL, REFERENCE	TITLE
Unpublished	Charakterization and repeat analysis of a freshwater pulfitefish Tetraodon nilotriviridis
2 (bases 1 to 656)	
Rossi-Collins, H. E., Jallou, O., Iustiva, C., Bouneau, L., Fisher, C. C.	

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**JOURNAL.** Unpublished  
**REFERENCE** 3 (bases 1 to 656)  
**AUTHORS** Genoscope.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (12-APR 2000) to The EMBL/Genbank/DBJ databases

**JOURNAL**  
submitting your  
**COMMENT**  
This sequence is a single read and was generated as part of a large scale clone-and-sequencing project of the *Helicobacter nigroviridis* genome. For more information, please take a look at <http://www.genoscope.cns.fr/Helicobacter>.

FEATURES	SOURCE	Location/Qualifiers
1. .656		
/outanism		"1st road on nigroviridis"
/th xref		"taxon: 9984"
/th xref		"11712"

BASE COUNT  
ORIGIN

	211	a	140	c	141	q	164	t
/curve - "basecount"	end : T7							
/curve -t t <sub>0</sub>								

Query Match: 7.8%; Score 66; DB 13; Length 656;  
Best local Similarity: 69.2%; Pred. No. 1,35-00;  
Microsites: 40; Indels 0; Gaps

[illegible][illegible]

Oy	745	gaclqlclqcl	754
		llll	ll
Db	124	gaclclclclcl	133



















```

APPLICATION NUMBER: US/08/091,569
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/750,081
FILING DATE: 27-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 16777/149/ALIE
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3695 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: si4_peptide
LOCATION: 17,...229
FEATURE:
NAME/KEY: mal_peptide
LOCATION: 230,...3118
FEATURE:
NAME/KEY: CDS
LOCATION: 176,...3118

```

[illegible]

```

SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/203,676
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/750,081
FILING DATE: 27-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 16777/149/ALLE
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3695 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 176..229
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 240..3118
FEATURE:
NAME/KEY: CDS
LOCATION: 176..3118

```

Search completed: March 4, 2002, 05:12:23  
Job time: 16997 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 4, 2002, 02:14:50 ; Search time 6206.58 Seconds  
(without alignments)  
1554.755 Million cell updates/sec

Title: US-08-956-991-8  
Perfect score: 898  
Sequence: 1 lgcgcgcgcgttcgaagcct.....gcgcacacatcgttcgtcga 898

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues  
22703874

( 1 number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*  
1: em\_estfun:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estom:\*  
5: em\_estpl:\*  
6: em\_estro:\*  
7: em\_estro:\*  
8: em\_estov:\*  
9: em\_hic:\*  
10: qd\_est1:\*  
11: qd\_est2:\*  
12: qd\_hic:\*  
13: qd\_gss:\*  
14: em\_gss\_fun:\*  
15: em\_gss\_hum:\*  
16: em\_gss\_inv:\*  
17: em\_gss\_pln:\*  
18: em\_gss\_pro:\*  
19: em\_gss\_rtd:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_other:\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Mat h	Length	DB ID	Description
C 1	224.4	25.0	279	10	BB336040 BB336040
C 2	223.8	24.9	291	10	BB333216 BB333216
C 3	218.4	24.3	252	10	BB381728 BB381728
C 4	214.8	23.7	270	10	BB336019 BB336019
C 5	212.8	23.7	296	10	BB335137 BB335137
C 6	212.8	23.7	355	10	BB461534 BB461534
C 7	204.4	22.7	247	10	BB330789 BB330789
C 8	204.4	22.7	248	10	BB329941 BB329941
C 9	192.4	21.4	257	10	BB330426 BB330426
C 10	190.6	21.2	303	10	BB331552 BB331552
C 11	189.2	21.1	304	10	BB280255 BB280255
C 12	189	21.0	286	10	BB333910 BB333910

C 13	186	20.7	499	11	B1185857
C 14	184.6	20.6	280	10	BB331669
C 15	180	20.0	252	10	BB334670
C 16	175.8	19.6	239	10	BB332000
C 17	174.4	19.4	272	10	BB330770
C 18	159.8	17.8	216	10	BB328887
C 19	146.2	16.3	219	10	BB342624
C 20	137.8	15.3	212	10	BB330864
C 21	128.4	14.3	194	10	BB331191
C 22	122.4	13.6	287	11	BF928859
C 23	117	13.0	163	10	BB330436
C 24	111.6	12.4	229	10	BB333024
C 25	100.8	11.2	1013	13	CNS02AK7
C 26	92.6	10.3	965	13	CNS02WH1
C 27	64.6	7.2	643	13	AZ790921
C 28	63.8	7.1	794	13	CNS03C7U
C 29	58	6.5	605	13	FR0021994
C 30	58	6.5	618	13	FR0022042
C 31	57.4	6.4	938	13	CNS03410
C 32	55.6	6.2	443	13	FR0022007
C 33	54.4	6.1	385	13	CNS04LEK
C 34	52.6	5.9	399	13	FR0022008
C 35	51	5.7	1101	13	CNS057K8
C 36	50	5.6	564	13	FR0030722
C 37	45	5.0	1101	13	CNS00KR2
C 38	44.4	4.9	73	13	CNS03K9Y
C 39	43.4	4.8	430	13	FR0022072
C 40	43	4.8	419	10	A1750970
C 41	41.8	4.7	405	10	A1005408
C 42	41.8	4.7	969	13	CNS05MTE
C 43	40.8	4.5	573	11	BF266074
C 44	39.6	4.4	930	10	A1566074
C 45	38.8	4.3	423	11	BC263635

#### ALIGNMENTS

RESULT 1  
BB336040/c 279 bp mRNA EST 11-Jul-2000  
LOCUS BB336040 RIKEN full-length enriched, 10 days neonate medulla  
DEFINITION Oblongata Mus musculus cDNA clone BB3046F24 3' similar to AF217525  
Homo Sapiens cDNA clone cDSC1 Down syndrome cell adhesion molecule  
(USCAM) mRNA, mRNA sequence.

ACCESSION BB336040  
VERSION BB336040.1 GI:9044803  
KEYWORDS EST.  
SOURCE house mouse  
ORGANISM Mus musculus

#### REFERENCE

Enkaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus. I (bases 1 to 279)  
Kono, H., Aizawa, K., Akahira, S., Akiyama, J., Arikawa, T., Carlini, P., Endo, T., Fukuda, S., Fukunishi, Y., Hata, A., Hayatsu, N., Hirozane, T., Horii, F., Ishii, Y., Ishikawa, J., Itoh, M., Izawa, M., Kadoya, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N., Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C., Kusakabe, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y., Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K., Shibata, Y., Shigemoto, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Tagawa, A., Takahashi, F., Tomioka, N., Toyota, T., Tsunoda, Y., Watabiki, A., Watanabe, S., Yamamura, T., Yamane, I., Yano, R., Yasunishi, A., Yokota, T., Yoshida, K., Yoshiki, A., Yoshino, M., Muramatsu, M., and Hayashizaki, Y.  
RIKEN Mouse ESTs (Kono, H., et al.)  
Unpublished (2000)

TITLE JOURNAL  
COMMENT  
Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222



Tel: 81-45-503-9222  
Fax: 81-45-503-9216

QY	447	caatgctcaatagccctgctgtagtgcctcgllgcacataaaccccatgctccactgagctgga	506
Db	247	CAATGCTCAATAGCCCTTGATGAGGGGGGTTGGCCATCAAAACCCCATGCTGCACCTGAGCGTGA	186
QY	507	tgtatgcagaccttggacatctcttgatctcacaatctcggagagatcctggaggtcttctgcact	566
Db	187	TGTATGCAGACCTTGACATCTT-TTGATCTTCATATCGGAGAGGATCTGAGGGGTCTTGAGACT	126
QY	567	gttatttgaattatttcacacgtccctcccgatattgaattgatatgcatggcagaagaagaa	626
Db	128	GTTATTGGAATTATTTCACACGCTCTCCCGATTGCAATTGATTCATGGCAGAGACAAAGAA	69
QY	627	ccggaatctctctctcactglttggcaaatctgcagcgtatagatatcaacttctctcccaac	688
Db	68	CCGGAATCTCTCTCAGCTGGGCTGAAATCTCAGCGTAGATATCATGTTCCCTCCACGC	9
QY	687	tctcttgat	694
Db	8	TCTCTGCT	1

























GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 4, 2002, 05:22:27 : Search time 578.64 Seconds  
(without alignments)  
1330.496 Million cell updates/sec

Title: US-08-956-991-8  
Perfect score: 898  
Sequence: 1 tgcgcgcgttgcagcct.....gcgccacatcgttgcctga 898

Scoring table: IDENTITY-NUC  
Gapop 10.0, Gapext 1.0

Searched: 930021 seqs, 428662619 residues  
1 number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N.Geneseq.1101.\*  
1: /SID52/gcgdata/geneseq/geneseq/NA1980.DAT.\*  
2: /SID52/gcgdata/geneseq/geneseq/NA1981.DAT.\*  
3: /SID52/gcgdata/geneseq/geneseq/NA1982.DAT.\*  
4: /SID52/gcgdata/geneseq/geneseq/NA1983.DAT.\*  
5: /SID52/gcgdata/geneseq/geneseq/NA1984.DAT.\*  
6: /SID52/gcgdata/geneseq/geneseq/NA1985.DAT.\*  
7: /SID52/gcgdata/geneseq/geneseq/NA1986.DAT.\*  
8: /SID52/gcgdata/geneseq/geneseq/NA1987.DAT.\*  
9: /SID52/gcgdata/geneseq/geneseq/NA1988.DAT.\*  
10: /SID52/gcgdata/geneseq/geneseq/NA1989.DAT.\*  
11: /SID52/gcgdata/geneseq/geneseq/NA1990.DAT.\*  
12: /SID52/gcgdata/geneseq/geneseq/NA1991.DAT.\*  
13: /SID52/gcgdata/geneseq/geneseq/NA1992.DAT.\*  
14: /SID52/gcgdata/geneseq/geneseq/NA1993.DAT.\*  
15: /SID52/gcgdata/geneseq/geneseq/NA1994.DAT.\*  
16: /SID52/gcgdata/geneseq/geneseq/NA1995.DAT.\*  
17: /SID52/gcgdata/geneseq/geneseq/NA1996.DAT.\*  
18: /SID52/gcgdata/geneseq/geneseq/NA1997.DAT.\*  
19: /SID52/gcgdata/geneseq/geneseq/NA1998.DAT.\*  
20: /SID52/gcgdata/geneseq/geneseq/NA1999.DAT.\*  
21: /SID52/gcgdata/geneseq/geneseq/NA2000.DAT.\*  
22: /SID52/gcgdata/geneseq/geneseq/NA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	898	100.0	898	19	AAV31986
2	502	55.9	6413	19	AAV31988
3	502	55.9	6604	19	AAV31981
4	244.8	27.3	5824	22	AA158196
5	43	4.8	5720	22	AA158196
6	38.8	4.3	1225	18	AA158196
7	38.8	4.3	1228	21	AA158196
8	38.8	4.3	1238	19	AAV48217
9	38.8	4.3	2793	21	AAV48217
10	38.8	4.3	2793	19	AAV48217
11	37.8	4.2	1648	13	AAO25532

C 12	37.2	4.1	1887	22	AAH25568
C 13	37.2	4.1	8269	22	AA164010
C 14	37	4.1	1341	20	AA198114
C 15	37	4.1	42235	20	AA198035
C 16	36.2	4.0	1308	18	AA160756
C 17	36.2	4.0	1308	19	AAV48215
C 18	36.2	4.0	1308	21	AA251042
C 19	35.4	3.9	690	20	AA198113
C 20	35.2	3.9	7702	21	AA198739
C 21	35.2	3.9	7702	21	AA291908
C 22	35.2	3.9	7702	21	AA259133
C 23	35.2	3.9	7705	22	AA198405
C 24	35.2	3.9	7705	22	AA198405
C 25	35.2	3.9	7741	22	AA198405
C 26	35	3.9	475	22	AA198405
C 27	34.6	3.9	213	22	AA152138
C 28	34.2	3.8	4134	21	AA238152
C 29	34	3.8	1307	21	AA163927
C 30	33.8	3.8	1032	22	AA163927
C 31	33.8	3.8	1032	21	AA163927
C 32	33.8	3.8	6253	19	AA163927
C 33	33.8	3.8	6254	19	AA163927
C 34	33.8	3.8	6254	22	AA163927
C 35	33.8	3.8	7193	19	AA163927
C 36	33.8	3.8	9436	15	AA163927
C 37	33.4	3.7	249	22	AA119369
C 38	33.4	3.7	249	22	AA144559
C 39	33.4	3.7	249	22	AA105056
C 40	33.4	3.7	456	22	AA114406
C 41	33.4	3.7	456	22	AA135779
C 42	33.4	3.7	456	22	AA104229
C 43	33.4	3.7	476	22	AA110073
C 44	33.4	3.7	476	22	AA113131
C 45	33.4	3.7	476	22	AA100016

## ALIGNMENTS

RESULT 1  
AAV31986 standard. cDNA: 898 BP.  
ID AAV31986  
XX AAV31986;  
XX 28-SEP-1998 (first entry)  
XX Mouse Down syndrome-cell adhesion molecule US-CAM mid cDNA.  
DE  
XX DS-CAM; Down syndrome-cell adhesion molecule; neural cell;  
XX signal transduction; trisomy 21; mental retardation;  
XX holoprosencephaly; corpus callosum agenesis;  
XX schizencephaly; diagnosis; assay; mouse; ds; ss.  
XX  
XX Mus sp.  
XX W09817795-A1.  
XX 30-APR-1998.  
XX 23-OCT-1997; 97WO-US19547.  
XX 25-OCT-1996; 96US-0029322.  
XX  
XX (CEDA-) CEDANS SIMAI MEDICAL CENT.  
XX Korenberg JR;  
XX WPI: 1998-271791/24.  
XX New isolated Down's Syndrome-cell adhesion molecule - used to  
XX develop products for detection, diagnosis and therapy of  
XX developmental and neurological abnormalities  
PT

Nucleotide sequenc  
Human polynucleoti  
Nucleotide sequenc  
Nucleotide sequenc  
Nucleotide sequenc  
Rat aortic prefere  
Rat aortic prefere  
Nucleotide sequenc  
Human protein tyro  
Human protein tyro  
LAR tyrosine phosph  
Human EST-derived  
Human cDNA encodin  
Human cDNA encodin  
Probe #7765 used t  
Probe #70824 used t  
Human NR-CAM gene  
DNA encoding prote  
Mouse sodium chann  
Streptomyces clavu  
Nucleotide sequenc  
Coding sequence of  
Human cDNA clone C  
Streptomyces clavu  
Blood transmitschl  
Probe #9302 for ge  
Probe #13245 used  
Probe #5087 used t  
Probe #4339 for ge  
Probe #4465 used t  
Probe #4220 used t  
Probe #6 for gene  
Probe #7 used to m  
Probe #7 used to m

XX  
PS      Claim 2; Page 81; 109pp; English.

This cDNA sequence comprises the middle region of a cDNA clone for murine Down syndrome-cell adhesion molecule (DS-CAM), a member of a novel subclass of the 19 superfamily with homology to neural cell adhesion molecules. The 5' region (see AAV31985) and 3' region (see AAV31987) of the clone are also provided. The murine DS-CAM clone was isolated from a C57 Black/6 mouse brain cDNA library using human DS-CAM cDNA (see AAV31981 and AAV31986) as probe. The invention also provides human DS-CAM proteins (see AAV42086-87), as well as expression vectors and host cells, transgenic animals, antisense oligonucleotides, and primers useful for amplification of DS-CAM nucleic acids. DS-CAM polypeptides are associated with developmental and neurological processes. They can be used in e.g. neural prosthetic devices used in entubulation methods of repairing (regenerating) damaged or severed peripheral nerves. The products can also be used in detection, diagnosis and therapy of developmental and neurological abnormalities such as Down syndrome, mental retardation, holoprosencephaly, agenesis of the corpus callosum, or schizophrenia. Antisense oligonucleotides are used for inhibiting translation of mRNA.

XX	Sequence	898 BP: 174 A: 222 C: 245 G: 257 T: 0 other:	
98	Query Match	100.0%; Score 898; DB 19; Length 898;	
99	Best Local Similarity	100.0%; Pred. No. 2,7e-273;	
100	Matches 898: Conservative	0; Mismatches 0; Indels 0; Gaps 0	
101	1	tgccgcgcggttcgaagccgtgtactacacagggccatctgcgttaattacaaatgcttgcacg 60	
102	1		60
103	1	tgccgcgcgcggtttgcgaagccgtgtactacacagggccatctgcgttaattacaaatgcttgcacg 60	
104	61	ggtgtatacaacttcgcgtctccgcgttggtgtcattacatctgatactgttgaacttggaaattccccc 120	
105	61		120
106	61	ggtgtatacaacttcgcgtctccgcgttggtgtcattacatctgatactgttgaacttggaaattccccc 120	
107	121	cgtctcgttactcccggtatgccttatctgtgtatgcgcgaatgattccgcgtctcttgtatactggtctt 180	
108	121		180
109	121	cgtctcgttactcccggtatgccttatctgtgtatgcgcgaatgattccgcgtctcttgtatactggtctt 180	
110	181	tggtagagcccttcacgaatgaacctgtgactctctgaatatagaagtgtggtcttcccaatgtgaattccc 240	
111	181		240
112	181	tggtagagcccttcacgaatgaacctgtgactctctgaatatagaagtgtggtcttcccaatgtgaattccc 240	
113	181		240
114	181	tggtagagcccttcacgaatgaacctgtgactctctgaatatagaagtgtggtcttcccaatgtgaattccc 240	
115	241	tggtagagcccttcacgaatgaacctgtgactctctgaatatagaagtgtggtcttcccaatgtgaattccc 300	
116	241		300
117	241	tggtagagcccttcacgaatgaacctgtgactctctgaatatagaagtgtggtcttcccaatgtgaattccc 300	
118	301	tctctgccaatccgcgtctcttgcgcgttcacatgcgagatgagctctgtatgcttgagagaaaggtctgatat 360	
119	301		360
120	301	tctctgccaatccgcgtctcttgcgcgttcacatgcgagatgagctctgtatgcttgagagaaaggtctgatat 360	
121	301		360
122	301	tctctgccaatccgcgtctcttgcgcgttcacatgcgagatgagctctgtatgcttgagagaaaggtctgatat 360	
123	361	caatgatacgtgtgcgtgcgttccacgtctgcaggtgcgagaaacattcttgcctctcctgcagccgaattccc 420	
124	361		420
125	361	caatgatacgtgtgcgtgcgttccacgtctgcaggtgcgagaaacattcttgcctctcctgcagccgaattccc 420	
126	421	acgcgtcctgtatattcttcttgcattacacactgtgtcatagacctcttgatagtggtgcgtctgtccatc 480	
127	421		480
128	421	acgcgtcctgtatattcttcttgcattacacactgtgtcatagacctcttgatagtggtgcgtctgtccatc 480	
129	481	aaaccccctgtctccaccctgcagcttgatagtgtgcgaagctcttgacacacacctgtgtcttcaatt 540	
130	481		540
131	481	aaaccccctgtctccaccctgcagcttgatagtgtgcgaagctcttgacacacacctgtgtcttcaatt 540	
132	541	cggtagagatcttcaggggtctctcttgcaatctgtgaaattatctccacacgcgtctcccccgtatg 600	
133	541		600
134	541	cggtagagatcttcaggggtctctcttgcaatctgtgaaattatctccacacgcgtctcccccgtatg 600	
135	601	aatttatatgcatgtgcagagaaagaaacacggaaatctctctctaacgcttggtcgaatatcttgcga 660	
136	601		660
137	601	aatttatatgcatgtgcagagaaagaaacacggaaatctctctctaacgcttggtcgaatatcttgcga 660	
138	601		660
139	601	aatttatatgcatgtgcagagaaagaaacacggaaatctctctctaacgcttggtcgaatatcttgcga 660	
140	601		660

OY	661	gcgagatcatcttcttcccaacctctggtgataacagatcaggccacttcaagg	720
Db	661	gcgtggtatcatcttcttcccaacctcttggtagatacagtaaggccacttcaagg	720
OY	721	ttaatgctctgctctctcttcccaagacacatgatagggctctcccaaggctgacg	780
Db	721	ttaatgctctgctctctcttcccaagacacatgatagggctctcccaaggctgacg	780
OY	781	ctcatctctctcttgaacctgagtcgacgggtggtgtggtatgaattatcatctgacg	840
Db	781	ctcatctctctcttgaacctgagtcgacgggtggtgtggtatgaattatcatctgacg	840
OY	841	gaattctctctggtggtctcatgacatcttgctgaacggtcgtctccacatcgttcgtga	898
Db	841	gaattctctctggtggtctcatgacatcttgctgaacggtcgtctccacatcgttcgtga	898
RESULT 2			
ID	AAV31988/c		
XX	AAV31988 standard; cDNA: 6413 bp.		
XX	AAV31988:		
XX	28-SEP-1998 (first entry)		
XX	Human Down syndrome-cell adhesion molecule DS-CAM2 cDNA.		
XX	DS-CAM2: Down syndrome-cell adhesion molecule; neural cell;		
KW	signal transduction; trisomy 21; mental retardation;		
KW	holoprosencephaly; corpus callosum agenesis;		
KW	schizencephaly; diagnosis; assay; human; ds; ss.		
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	453..5168	
FT		/ftatg a	
XX	W09817795-A1.		
XX	30-APR-1998.		
PD	23-OCT-1997: 97NW-0519547.		
PF	25-OCT-1996: 96WS-0029422.		
XX			
XX	(CHDA-) CHDANKS SIMAI MPMOVAL, CENT.		
XX	Korenberg JR:		
DR	WPI: 1998-271791/24.		
DR	P-PSDB: AAM42087.		
PT	New isolated Down's Syndrome-cell adhesion molecule - used to		
PT	develop products for detection, diagnosis and therapy of		
XX	developmental and neurological abnormalities		
XX	Claim 3: Page 83-90; 109pp; English.		
<p>This cDNA clone codes for Down syndrome-cell adhesion molecule DS-CAM2 (see AAM42087), an extracellular, soluble protein belonging to a novel subclass of the Ig superfamily with highest homology to neural cell adhesion molecules. A modified direct cDNA selection technique was applied to bacterial and P1 artificial chromosomes between EHS2 and MX1 by using cDNA from trisomy 21 human foetal brain. A unique cDNA fragment, designated E51 (see AAV31982), was obtained and used to screen a trisomy 21 human foetal brain cDNA library. Further clones were obtained by exon trapping, and the complete DS-CAM2 cDNA sequence was identified. DS-CAM2 is a splice variant of membrane-bound DS-CAM1 (see AAM42086 and is deleted for the entire transmembrane domain. The DS-CAM gene spans 900-1200 kb of genomic DNA and maps p1 chromosome</p>			





RESULT	/5
AA#97891/C	
ID	AA#97891 standard; cDNA; 5720 bp.
XX	
AC	AA#97891;
XX	
DT	01-JUN-2001 (first entry)
XX	
DE	Human secreted protein cDNA, SEQ ID NO: 18.
XX	
KW	Human; secreted protein; immunomodulatory; antiproliferative

XX	16-OCT-1997	(first entry)
DT		
XX	Human aortic	preferentially expressed gene 1
DE		

```

XX Human: aortic preferentially expressed gene; APEG-1; aorta;
KM specific; inhibition; growth; smooth; muscle; cell; vascular;
KM proliferation; artery; restenosis; atherosclerosis; transplant;
KM diabetes; ds.
OS Homo sapiens.
XX Key Location/Qualifiers
FH CDS 78..419
FT FT /*lag-a
      /label= APEG-1
XX W09700953-A1.
PD 09-JAN-1997.
XX 20-JUN-1996; 96WO-US10636.
PF 22-JUN-1995; 95US-0494577.
F (HARD ) HARVARD COLLEGE.
X-
XX Hsieh C, Lee M:
PI MPI: 1997-087380/08.
XX P-PSUB; AAU10566.
DR Aortic preferentially-expressed gene APEG-1 - useful to develop
PT prods. for diagnosis and treatment of vascular conditions, e.g.
XX PT restenosis, atherosclerosis and diabetes
XX Claim 10; Pages 37-38; 58pp; English.
XX The present sequence encodes the human aortic preferentially
CC expressed gene 1 (APEG-1) polypeptide, which is aorta specific and
CC inhibits the growth of smooth muscle cells. The APEG-1 polypeptide
CC can be used to inhibit vascular smooth muscle cell proliferation
CC In an animal via its introduction into an artery of the animal.
CC Products derived from APEG-1 and the APEG-1 polypeptide can be
CC used to diagnose and treat vascular conditions, e.g., restenosis,
CC atherosclerosis, transplant arteriosclerosis or diabetes.
CC Rat APEG-1 DNA, identified by differential expression in the aorta
CC compared to other organs, was used to obtain human APEG-1 DNA.
XX SO Sequence 1225 BP; 249 A; 363 C; 382 G; 231 T; 0 other:
(
    try Match          4.3%; Score 38.8; DB 18; Length 1225;
    st Local Similarity 62.28; Pred. No. 0.056;
    Matches 61; Conservative 0; Mismatches 37; Indels 0; Gaps
OY 592 ccccatgaatgatagatgacgtgcagagaagaaccggaaattcttcactgttgga 65]
DB 378 CACCATTAATCTATGCGCGCTTGCAAGTGTGAAGAACCAGCATGGCCAGCTTCGACGCCA 319
OY 652 aaatcgcagcgtagatacttcccctccccacctc 689
DB 318 GGATCCGCAGC CGGCACAGCCACCCCTCAGCCTCTCC 381
RESULT 7
AAZ51043/C
ID AAZ51043 standard; CDNA: 1228 BP.
XX
AC AAZ51043;
XX
DE 05-JUN-2000 (first entry)
XX Human aortic-preferentially-expressed gene-1 cDNA.
XX Human: aortic-preferentially-expressed gene-1; APEG-1;
KM aorta; striated muscle cell; vascular smooth muscle cell; VSMC;
```

KM		antiartherosclerotic; vasotropic; cis-acting transcriptional repressor;
KW		treatment; diagnosis; vascular disease; atherosclerosis; restenosis;
RN		chromosome 2q33-34; ss.
XX		
OS	Homo sapiens.	
XX		
FH	key	Location/Qualifiers
FT	CDS	78..419
FT		/*Lay- d
PN		/product = "APEG-1 protein"
XX		
MO	WO200009689-A2.	
XX		
PD	24-FEB-2000.	
XX		
PF	11-MAY-1999;	99WO-US10298.
XX		
PR	14-AUG-1998;	98US-0134250.
PR	30-APR-1999;	99US-0303069.
XX		
PA	(HARD ) HARVARD COLLEGE.	
PI		
PI	Lee M, Hsieh C;	
DR	WPI: 2000-224334/19.	
DR	P-PDB: AAT70077.	
PT		
PT	New gene useful for treating and diagnosing vascular diseases comprises	
PT	a single gene encoding aortic-specific and striated-specific muscle	
PT	cell isoforms -	
PS		
PS	Disclosure: Fig 16: 88pp; English.	
XX		
CC	The present sequence is a cDNA encoding aortic-preferentially-expressed	
CC	gene-1 (APEG-1) protein from a human lambda gt10 aortic 5'-stretch cDNA	
CC	library. APEG-1 gene, located on human chromosome 2q33-34,	
CC	encodes two muscle cell protein isoforms, one specific to aortic smooth	
CC	muscle cells designated APEG-1 protein and the other specific to	
CC	striated muscle cells designated SPEG protein.	
CC	APEG-1 protein can be administered to vascular smooth muscle cells	
CC	(VSMC) to inhibit their proliferation or migration at the site of	
CC	vascular injury. The present sequence is used to screen VSMC-specific	
CC	enhancer sequence which is used to direct VSMC-specific expression. A	
CC	cis-acting transcriptional repressor sequence found in the 5' region of	
CC	APEG-1 gene is useful to detect compounds that bind to the repressor and	
CC	increase APEG-1 expression in VSMC. APEG-1 is useful for treating and	
CC	diagnosing vascular diseases such as atherosclerosis and restenosis.	
XX		
SQ	Sequence 1228 BP; 249 A; 363 C; 382 G; 231 T; 3 other;	
	Query Match	4.3%; Score 38.0; DB 21; Length 1228;
	Best Local Similarity	62.2%; Pred. No. 0.056; 37; Indels 0; Gaps 0;
	Matches 61; Conservative 0; Mismatches	
QY	592 CCCGTAAGATTGTATCATCATCGACGAGAAGAAACAACCAGCACTTCTCACTGTGGCA 651	
Db	378 CACCATACTACATTTGCAGCCGCTTTGCCAAGTGTCGAACAACAGCATGCCACGCTTCGACGCCA 319	
QY	652 aaatcgtacagcgtagatalcatctccctcccactcc 689	
Db	318 GGATCCGACGACCGGACAGACGCCCACCCCTCAGCTCCTCC 281	
	RESULT 8	
ID	AAVA48217/c	
AC	AAVA48217 standard; DNA; 1238 bp.	
XX		
AC	AAVA48217;	
XX		
DT	09-NOV-1998 (first entry)	
XX		
DE	Human aortic-preferentially-expressed gene-1.	

```

XX ds: human; aortic preferentially expressed protein 1; smooth muscle;
KM cell proliferation; developmental stage; tissue plasminogen activator;
XX p21 cell cycle; nitric oxide synthetase; gamma-interferon.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 88..429
XX /tag="a" "Aortic-preferentially-expressed protein-1"
XX /product="
XX /note="No start codon given"
XX
XX W09835040-A2.
XX
XX 13-AUG-1998.
XX
XX 06-FEB-1998; 98MO-US02441.
XX
XX 06-FEB-1997; 97US-0795868.
XX
XX (HARD ) HARVARD COLLEGE.
XX
XX Hsieh C, Lee M.
XX
XX WPI: 1998-447237/38.
XX
XX P-PSDB: AAW77047.
XX
XX Novel human, rat or mouse aorta or striated-muscle preferentially
XX expressed proteins - useful for treating e.g. atherosclerosis
XX
XX Disclosure: Fig 16; 88pp; English.
XX
XX PS The aortic preferentially expressed protein 1 (APEG-1) can be used to
XX CC derive an enhancer/promoter. This linked to a polypeptide coding sequence
XX CC which regulates smooth muscle cell-specific expression of the polypeptide
XX CC coding sequence can be used as a method of inhibiting vascular smooth
XX CC muscle cell proliferation. The nucleic acids are used to direct
XX CC developmental stage-specific expression of a heterologous polypeptide
XX CC which is especially selected from tissue plasminogen activator (tPA),
XX CC p21 cell cycle inhibitor, nitric oxide synthetase, gamma-interferon,
XX CC atrial natriuretic proteins. These are used to inhibit the proliferation
XX CC of smooth muscle cells, e.g. for the treatment of atherosclerosis.
XX
XX SO Sequence 1238 BP; 251 A; 365 C; 387 G; 232 T; 3 other;
XX
XX Query Match 4.18; Score 38.8; DB 19; Length 1238;
XX Best Local Similarity 62.2%; Pred. No. 0.056;
XX Matches 61; Conservative 0; Mismatches 37; Indels 0; Gaps 0;
XX
XX OY 592 cccgcatgaatgtagatgcatgagagaagaaacggaaatctctctcaactgttgca 651
XX 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
XX DB 388 CACCATACCTACATGACCGCTTGTGCAAGTGTAGAAACACGATCGCCAGCTCTGCAGCCA 329
XX
XX OY 652 aaatctgacgctagatatacactctctctccaccctcc 689
XX 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
XX DB 328 GGATCCGACGCGGACGACGCCACCTCAGCTCCTCTCC 291
XX
XX RESULT 9
XX AAV48218/C
XX ID AAV48218 standard; DNA: 2793 BP.
XX
XX AC AAV48218;
XX
XX 09-NOV-1998 (first entry)
XX
XX Human striated muscle preferentially expressed gene.
XX
XX ds: human; striated muscle preferentially expressed protein;
XX smooth muscle; cell proliferation; developmental stage; gamma-interferon;
XX tissue plasminogen activator; p21 cell cycle; nitric oxide synthetase.

```

```

XX OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 3..1986
XX /tag="a" "Striated muscle preferentially expressed
XX /product="
XX /note="No start codon given"
XX
XX W09835040-A2.
XX
XX 13-AUG-1998.
XX
XX 06-FEB-1998; 98MO-US02441.
XX
XX 06-FEB-1997; 97US-0795868.
XX
XX (HARD ) HARVARD COLLEGE.
XX
XX Hsieh C, Lee M.
XX
XX WPI: 1998-447237/38.
XX
XX P-PSDB: AAW77048.
XX
XX Novel human, rat or mouse aorta or striated-muscle preferentially
XX expressed proteins - useful for treating e.g. atherosclerosis
XX
XX Claim 14; Fig 20; 88pp; English.
XX
XX PS The striated muscle preferentially expressed protein (SPEG) can be linked
XX CC to an enhancer/promoter derived from an aortic preferentially expressed
XX CC gene to regulate smooth muscle cell-specific expression. This can
XX CC be used as a method of inhibiting vascular smooth muscle cell
XX CC proliferation. The nucleic acids are used to direct developmental
XX CC stage-specific expression of a heterologous polypeptide which is
XX CC especially selected from tissue plasminogen activator (tPA), p21 cell
XX CC cycle inhibitor, nitric oxide synthetase, gamma-interferon, atrial
XX CC natriuretic proteins. These are used to inhibit the proliferation of
XX CC smooth muscle cells, e.g. for the treatment of atherosclerosis.
XX
XX SO Sequence 2793 BP; 499 A; 963 C; 920 G; 411 T; 0 other;
XX
XX Query Match 4.18; Score 38.8; DB 19; Length 2793;
XX Best Local Similarity 62.2%; Pred. No. 0.085;
XX Matches 61; Conservative 0; Mismatches 37; Indels 0; Gaps 0;
XX
XX OY 592 cccgcatgaatgtagatgcatgagagaagaaacggaaatctctctcaactgttgca 651
XX 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
XX DB 1947 CACCATACCTACATGACCGCTTGTGCAAGTGTAGAAACACGATCGCCAGCTCTGCAGCCA 1888
XX
XX OY 652 aaatctgacgctagatatacactctctctccaccctcc 689
XX 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
XX DB 1887 GGATCCGACGCGGACGACGCCACCTCAGCTCCTCTCC 1850
XX
XX RESULT 10
XX AA251044/C
XX ID AA251044 standard; cDNA: 2793 BP.
XX
XX AC AA251044;
XX
XX 05-JUN-2000 (first entry)
XX
XX Human striated muscle preferentially expressed gene partial cDNA.
XX
XX ds: human; striated muscle preferentially expressed gene; SPEG;
XX aortic-preferentially-expressed gene-1; APEG-1; striated muscle cell;
XX aorta; smooth muscle cell; antiatherosclerotic; vasodilator;
XX treatment; diagnosis; vascular disease; atherosclerosis; restenosis;
XX chromosome 2p33-34; ss.
XX

```



```

OS Homo sapiens.
XX Key Location/Qualifiers
FH 3..1988
FT CDS /ftag= a
Fv /product= "SPEG protein"
Fv /partial
PN M020009689-A2.
PD 24-FEB-2000.
XX 11-MAY-1999; 99MO-US10298.
XX 14-AUG-1998; 98US-0134250.
PR 30-APR-1999; 99US-0303069.
XX (HARR) HARVARD COLLEGE.
F Lee M, Hsieh C;
L.. WPI: 2000-22A334/19.
P-PDB: AAT70078.
XX New gene useful for treating and diagnosing vascular diseases comprises
PT a single gene encoding aortic-specific and striated-specific muscle
PT cell isoforms -
XX Claim 29: Fig 20; 88pp; English.
PS The present sequence is a partial cDNA encoding human striated muscle
CC preferentially expressed gene (SPEG) protein. This protein
CC was detected in striated muscles of skeletal and cardiac tissues.
CC It is one of the two muscle cell protein isoforms encoded by
CC the aortic-preferentially-expressed gene-1 (APEG-1) located on human
CC chromosome 2q33-34. The other protein isoform is specific to aortic
CC smooth muscle cells designated APEG-1 protein. The present sequence can
CC be used to identify striated muscle cell specific promoter which may be
CC useful to direct gene expression in striated muscle cells to treat
CC diseases associated with these muscles. APEG-1 gene is useful for
CC treating and diagnosing vascular diseases such as atherosclerosis and
CC restenosis.
SO Sequence 2793 BP; 499 A; 963 C; 920 G; 411 T; 0 other;

Query Match 4.3%; Score 38.8; DR 21; Length 2793;
      nest Local Similarity 62.2%; Pred. No. 0.085;
      ches 61; Conservative 0; Mismatches 37; Indels 0; Gaps

QY 592 ccccgatgaatgataagacgcaggagaagaaccggaaattcttctactgtttgca 651
    1111111111111111111111111111111111111111111111111111111
DB 1947 CACCATTACTCATTGTAGCCCCCTTTGGCAAGTGTAATAAACAGCATGCCGCATCGTCCTCACGCCA 1888

QY 652 aaatcttgacgcytatgatatacattcttcctccaccucc 689
    1111111111111111111111111111111111111111111111111111111
DB 1887 GGATTCGGACAGCGGACACACCACCCACTGACCTCTCC 1850

RESULT 11
AA025532/c
ID AA025532 standard; DNA; 1648 BP.
XX AA025532:
AC
XX
DT 22-NOV-1992 (first entry)
XX
DE Sequence of genomic clone conty. the entire Histidine-rich
DE protein (HisRP) gene.
XX
KW Malaria vaccine; histidine-rich protein; cytoadherence; ss.
OS Plasmodium lophurae.
```

XX	Key	location/Qualifiers
FH	sig_peptide	219..359 /*lag- a
FT	CDS	491..562 /*lag- b
FT	Intron	/product- propeptide 360..490
FT	misc_feature	/*lag- c 60..64
FT	misc_feature	/*lag- d /label- splice sequence 485..490
FT	misc_feature	/*lag- e /label= splice sequence 563..1477
FT	mat_peptide	/*lag- f /lag- i
PX		
PN	USJ116965-A.	
XX		
PD	25-MAY-1992.	
XX		
PF	26-AUG-1986;	86US-0900401.
PR	26-AUG-1986;	86US-0900401.
PA	(SLOK ) SLOAN KETTERING INST. CANCER. Pologe L, Kavelch JV; WP1; 1992-199590/24. P-PSDB; AAR24393.	
XX	Histidine-rich protein associated with Plasmodium knob phenotypic	
PT	and DNA encoding it, used for in vitro diagnosis of P.	
PT	Falciparum infection.	
XX		
PS	Disclosure; Fig 7A-B; Z';; English.	
XX		
XX	Two variants of HSRP are produced by P. falciparum. One is	
CC	associated with what is referred to as "knobby phenotype" (K30) and	
CC	"knobless phenotype" (K-). The "knobby" and "knobless" phenotypes	
CC	have been implicated in cytoadherence, which is characteristic of	
CC	erythrocyte infection. It has now been found that cDNA expressing	
CC	both K+ and K- HSRP can be obtained by the use of P. falciparum HSRP	
CC	expressing DNA. The genomic clone (AAU25512) is encoded from the	
CC	exons, separated by the signal poy. coding sequence from the	
CC	pro-sequence, confirming that synthesis of the protein occurs via	
CC	the pre-protein. Of two probes synthesised to the signal	
CC	peptide-encoding exon reveal multiple homologous DNA sequences in	
CC	the P. falciparum genome. The sequence of mature proteins is arranged	
CC	in numerous tandem repeats with up to nine histidine residues in a	
CC	row, similar to other plasmodium proteins for which sequence data	
CC	have so far been reported.	
XX		
SQ	Sequence 1648 BP; 569 A; 488 C; 150 G; 441 T; 0 other;	
	Query Match	A.2%; Score 37.8; DB 13; Length 1648;
	Best local Similarity	58.4%; Pred. No. 0.13;
	Matches	66; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
OY	267 lccgcgtagagtgatcctcgcttcgtggcctacccttgcacatcgacttcgtgcgaac	326
DB	1390 TGCAATGCTCATGTAGTGAATGGTGTTGTCATGTCATCGTAGTGCGTGTGGTGC	1331
OY	327 atcgcatgctgtcagatlgagaagaaggatgatatacaatcagtgtgcgaatc	379
DB	1330 ATCTGATGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1278
RESULT	12	
	AAH25568/c	







Oy 557 ttcttgacactgagatt 573  
| | | | | | | |  
Db 21611 TCCACGTCCTCTACGGT 21595

Search completed: March 4, 2002, 05:23:50  
Job time: 17659 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 4, 2002, 05:12:23 ; Search time 226.99 seconds  
(without alignments) 895.974 Million cell updates/sec

Title: US-08-956-991-8

Perfect score: 898  
Sequence: 1 tgcgcgcgcgttcgaagct.....gcgcacacatcgtcgtga 898

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 351'03 seqs, 11323899 residues

1 number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Issued\_patents\_NA:\*  
1: /cgn2\_6/p/odata/2/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/p/odata/2/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/p/odata/2/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/p/odata/2/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/p/odata/2/ina/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/p/odata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
C 1	38.8	4.3	1225	1	US-08-494-577-11 Sequence 11, Appl
C 2	38.8	4.3	1238	2	US-08-795-868-11 Sequence 11, Appl
C 3	38.8	4.3	2793	2	US-08-795-868-13 Sequence 13, Appl
C 4	36.4	4.1	7218	1	US-08-232-463-14 Sequence 14, Appl
C 5	36.2	4.0	342	1	US-08-494-577-2 Sequence 2, Appl
C 6	36.2	4.0	1308	2	US-08-494-577-2 Sequence 2, Appl
C 7	36.2	4.0	1308	2	US-08-494-577-1 Sequence 1, Appl
C 8	36.2	4.0	1308	2	US-08-494-577-1 Sequence 1, Appl
C 9	31.8	3.5	1755	1	US-08-068-395A-2 Sequence 1, Appl
C 10	31.8	3.5	1755	1	US-08-464-365-2 Sequence 1, Appl
C 11	31.4	3.5	1557	3	US-09-043-123-1 Sequence 2, Appl
C 12	31	3.5	1632	1	US-08-324-243-34 Sequence 34, Appl
C 13	31	3.5	1632	1	US-08-532-390-34 Sequence 34, Appl
C 14	31	3.5	1632	1	US-08-717-294-34 Sequence 34, Appl
C 15	31	3.5	1632	1	US-08-717-294-34 Sequence 34, Appl
C 16	31	3.5	2481	1	US-08-324-243-35 Sequence 35, Appl
C 17	31	3.5	2481	1	US-08-324-243-35 Sequence 35, Appl
C 18	31	3.5	2481	1	US-08-532-390-35 Sequence 35, Appl
C 19	31	3.5	2481	1	US-08-532-390-35 Sequence 35, Appl
C 20	31	3.5	2481	1	US-08-532-390-35 Sequence 35, Appl
C 21	30.8	3.4	2870	2	US-08-468-036-28 Sequence 28, Appl
C 22	30.8	3.4	2870	2	US-08-376-843-28 Sequence 28, Appl
C 23	30.8	3.4	4057	4	US-08-894-997-49 Sequence 49, Appl
C 24	30.6	3.4	3549	4	US-09-008-097-5 Sequence 5, Appl
C 25	30	3.3	3568	4	US-09-103-5108-1 Sequence 1, Appl
C 26	29.8	3.3	375	3	US-09-167-354-8 Sequence 8, Appl
C 27	29.8	3.3	1335	2	US-08-985-090-3 Sequence 3, Appl

C 28	29.8	3.3	1335	3	US-09-165-543-3 Sequence 3, Appl
C 29	29.8	3.3	1335	3	US-09-167-354-6 Sequence 6, Appl
C 30	29.8	3.3	2614	2	US-08-795-868-15 Sequence 15, Appl
C 31	29.8	3.3	2689	2	US-08-985-090-1 Sequence 1, Appl
C 32	29.8	3.3	2689	2	US-08-985-090-1 Sequence 1, Appl
C 33	29.8	3.3	2699	3	US-09-167-354-5 Sequence 5, Appl
C 34	29.6	3.3	206	4	US-09-476-124-11 Sequence 11, Appl
C 35	29.6	3.3	206	4	US-09-476-124-43 Sequence 43, Appl
C 36	29.6	3.3	1013	4	US-09-475-316A-30 Sequence 30, Appl
C 37	29.6	3.3	2424	4	US-09-234-393-1 Sequence 1, Appl
C 38	29.6	3.3	2525	4	US-09-234-393-39 Sequence 39, Appl
C 39	29.6	3.3	2528	4	US-09-234-393-37 Sequence 37, Appl
C 40	29.6	3.3	2528	4	US-09-234-393-41 Sequence 41, Appl
C 41	29.6	3.3	2571	4	US-09-234-393-12 Sequence 12, Appl
C 42	29.6	3.3	2774	4	US-09-363-189B-5 Sequence 5, Appl
C 43	29.6	3.3	4285	4	US-09-040-774-1 Sequence 1, Appl
C 44	29.4	3.3	890	4	US-09-475-316A-34 Sequence 34, Appl
C 45	29.4	3.3	913	4	US-09-475-316A-32 Sequence 32, Appl

#### ALIGNMENTS

RESULT 1  
US-08-494-577-11/c  
Sequence 11, Application US/08494577  
Patent No. 5786171  
GENERAL INFORMATION:  
APPLICANT: Lee, Mn-Bn  
TITLE OF INVENTION: Hsieh, Chung-Ming  
TITLE OF INVENTION: ADRICT: PREFERENTIALLY EXPRESSED GENE AND  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESSES:  
ADDRESS: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER RELEVABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/494,577  
FILING DATE: 22 JUN 1995  
CLASSIFICATION: 436  
ATTORNEY/AGENT INFORMATION:  
NAME: Fraser, Janis K.  
REGISTRATION NUMBER: 34,819  
REFERENCE/DOCKET NUMBER: 05434/012001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617/542-5070  
TELEFAX: 617/542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1225 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
US-08-494-577-11

Query Match 4.3%; Score 38.8; DB 1; Length 1225;  
Best Local Similarity 62.2%; Pred. No. 0.0056;  
Matches 61; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

592 cccgatalgaatgatgacatgagagagaaacccggaattcttcacactgttgca 651  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||



```
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232.463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935.313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 1MMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZgpt-F15
US-08-232-463-14

Query Match          4.18; Score 36.4; DB 1; Length 7218;
Best Local Similarity 11.6%; Pred. No. 0.1; Mismatches 156; Indels 0; Gaps 0;
Matches 40; Conservative 150;

551 tgggggttcctgcacgtgagtgatattatccacgagctccctccgtatgaattgtagc 610
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 979 TCGGCTCTACTACTATTTTCTCTGCTTGCCATACGCTCACAGATTAATTCGAGC 1038
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 611 atgcagagagaagaacccggaattctctctacacgttgcaaaatctcgagcgtagatat 670
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1039 TTGGCTCGACGTCGAGGAGCTTCGATATATATATATATATATATATATATATATAT 1098
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 671 cacttcctccaccctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 730
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1099 YYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYY 1158
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 731 tctctctctctccagcgagacaalggcctcctccagcctcctcctcctcctcctcctcct 790
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1159 YYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYY 1218
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 791 cctttacacgtcagcagcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 850
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1219 YYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYY 1278
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 851 gtagacatcagacgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 896
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1279 YYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYY 1324
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 5
US-08-494-577-2/C
Sequence 2, Application US/08494577
Patent No. 5786171
GENERAL INFORMATION:
APPLICANT: Lee, Mu-En
APPLICANT: Hsieh, Chung-Ming
TITLE OF INVENTION: AORTIC PREFERENTIALLY EXPRESSED GENE AND
TITLE OF INVENTION: USES THEREOF
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
```

```
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/494.577
FILING DATE: 22-JUN-1995
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Fraser, Janis K.
REGISTRATION NUMBER: 34,819
REFERENCE/DOCKET NUMBER: 05433/012001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 342 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-494-577-2

Query Match          4.08; Score 36.2; DB 1; Length 342;
Best Local Similarity 60.8%; Pred. No. 0.019;
Matches 59; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 592 ccccgatgaatgatagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 651
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 301 CGCATATTCGTTGACCGGCTTGCACAGTGTAGAACCGCATCTCCCTCTCGAGGCCA 242
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 652 aaatctgacgcagatagatcctcctcctcctcctcctcctcctcctcctcctcctcct 688
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 241 CGATCTCTACTACTATTTTCTCTGCTTGCCATACGCTCACAGATTAATTCGAGC 205
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 6
US-08-795-868-2/C
Sequence 2, Application US/08795868
Patent No. 5846773
GENERAL INFORMATION:
APPLICANT: Lee, Mu-En
APPLICANT: Hsieh, Chung-Ming
TITLE OF INVENTION: A SINGLE GENE ENCODING AORTIC-SPECIFIC
TITLE OF INVENTION: AND STRIATED-SPECIFIC MUSCLE CELL ISOFORMS AND USES THEREOF
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson, P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/795.868
FILING DATE: 06-FEB-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
```



APPLICATION NUMBER: 08/494,577  
FILING DATE: 22-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Fraser, Janis K.  
REGISTRATION NUMBER: 34,819  
REFERENCE/DOCKET NUMBER: 05433/032001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-542-5070  
TELEFAX: 617-542-8906  
INFORMATION FOR SEQ. ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 342 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
US-08-795-868-2

Query Match 4.0%; Score 36.2; DB 2; Length 342;  
atches 59; Conservatively 0; Mismatches 18; Indels 0; Gaps 0;  
OY 592 ccccgatgaatgtagatgagcagagagaacgga tcttctcactgttgc 651  
DB 301 ccccatattcgttgaccgcttgacagctgtagaagacagcatctccctcagcagcca 242  
OY 652 aaatctgacgtagatatactctctccacctc 688  
DB 241 ggatcctcaaccgacagaccacccctcgccctcctc 205

RESULT 7  
US-08-494-577-1/c  
Sequence 1, Application US/08494577  
Patent No. 5786171  
GENERAL INFORMATION:  
APPLICANT: Lee, Mu-En  
APPLICANT: Hsieh, Chung-Ming  
TITLE OF INVENTION: AORTIC PREFERENTIALLY EXPRESSED GENE AND  
TITLE OF INVENTION: USES THEREOF  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/494,577  
FILING DATE: 22-JUN-1995  
CLASSIFICATION: 436  
ATTORNEY/AGENT INFORMATION:  
NAME: Fraser, Janis K.  
REGISTRATION NUMBER: 34,819  
REFERENCE/DOCKET NUMBER: 05433/012001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617/542-5070  
TELEFAX: 617/542-8906  
INFORMATION FOR SEQ. ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1308 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA

US-08-494-577-1

Query Match 4.0%; Score 36.2; DB 1; Length 1308;  
Best Local Similarity 60.8%; Pred. No. 0.043; 38; Indels 0; Gaps 0;  
Matches 59; Conservatively 0; Mismatches 38;

OY 592 ccccgatgaatgtagatgagcagagagaacggaatcttctcactgttgc 651  
DB 470 ccccatattcgttgaccgcttgacagctgtagaagacagcatctccctcagcagcca 411  
OY 652 aaatctgacgtagatatactctctccacctc 688  
DB 410 ggatcctcaaccgacagaccacccctcgccctcctc 374

RESULT 8  
US-08-795-868-1/c  
Sequence 1, Application US/08795868  
Patent No. 5846773  
GENERAL INFORMATION:  
APPLICANT: Lee, Mu-En  
APPLICANT: Hsieh, Chung-Ming  
TITLE OF INVENTION: A SINGLE GENE ENCODING AORTIC-SPECIFIC  
TITLE OF INVENTION: AND RELATED-SPECIFIC MUSCLE CELL ISOFORMS AND USES THEREOF  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson, P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/795,868  
FILING DATE: 06-FEB-1997  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/494,577  
FILING DATE: 22-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Fraser, Janis K.  
REGISTRATION NUMBER: 34,819  
REFERENCE/DOCKET NUMBER: 05433/012001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-542-5070  
TELEFAX: 617-542-8906  
INFORMATION FOR SEQ. ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1308 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
US-08-795-868-1

Query Match 4.0%; Score 36.2; DB 2; Length 1308;  
Best Local Similarity 60.8%; Pred. No. 0.043; 38; Indels 0; Gaps 0;  
Matches 59; Conservatively 0; Mismatches 38;

OY 592 ccccgatgaatgtagatgagcagagagaacggaatcttctcactgttgc 651  
DB 470 ccccatattcgttgaccgcttgacagctgtagaagacagcatctccctcagcagcca 411  
OY 652 aaatctgacgtagatatactctctccacctc 688  
DB 410 ggatcctcaaccgacagaccacccctcgccctcctc 374

[illegible]

1 GENBANK: U00001  
 2 APPLICANT: SEED, BRIAN  
 3 TITLE OF INVENTION: High level Expression of Proteins  
 4 NUMBER OF SEQUENCES: 40  
 5 CORRESPONDENCE ADDRESS:  
 6 ADDRESSEE: Fish & Richardson P.C.  
 7 STREET: 225 Franklin Street  
 8 CITY: Boston  
 9 STATE: Massachusetts  
 10 COUNTRY: U.S.A.  
 11 ZIP: 02110-2804  
 12 COMPUTER READABLE FORM:  
 13 MEDIUM TYPE: Floppy disk  
 14 COMPUTER: IBM PC compatible

```

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30B
CURRENT APPLICATION NUMBER: US/08/532.390
FILING DATE: 22-SEP-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/324,243
FILING DATE: 19-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: LECH, KAREN F.
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/294001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 1632 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-532-390-34

```

```

Query Match 3.5%; Score 31; DB 1; Length 1632;
Best Local Similarity 53.8%; Pred. No. 2.7;
Matches 64; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

```

```

OY 263 ctacacgcgcgtatgagtcgtctgctgagcctccttcgcaatccggtcttgc 322
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 555 ctccgcgcgcgtatgagtcgtctgctgagcctccttcgcaatccggtcttgc 496
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 323 gtacacgcgcgtatgagtcgtctgctgagcctccttcgcaatccggtcttgc 381
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 495 gttgcgtctgctgagtcgtctgctgagcctccttcgcaatccggtcttgc 437
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

RESULT 14
US-08-717-294-34/C
Sequence 34, Application US/08717294
Patent No. 6114148
GENERAL INFORMATION:
APPLICANT: SEED, BRIAN
TITLE OF INVENTION: HIGH LEVEL EXPRESSION OF
PROTEINS
NUMBER OF SEQUENCES: 110
CORRESPONDENCE ADDRESS:
ADDRESS: 176 Federal Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/717,294
FILING DATE: 20-SEP-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Elbing, Karen L
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/345001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200

```

```

TELEFAX: 617-428-7045
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 1632 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-717-294-34

```

```

Query Match 3.5%; Score 31; DB 5; Length 1632;
Best Local Similarity 53.8%; Pred. No. 2.7;
Matches 64; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

```

```

OY 263 ctacacgcgcgtatgagtcgtctgctgagcctccttcgcaatccggtcttgc 322
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 555 ctccgcgcgcgtatgagtcgtctgctgagcctccttcgcaatccggtcttgc 496
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 323 gtacacgcgcgtatgagtcgtctgctgagcctccttcgcaatccggtcttgc 381
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 495 gttgcgtctgctgagtcgtctgctgagcctccttcgcaatccggtcttgc 437
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

RESULT 15
PCT-US95-11511-34/C
Sequence 34, Application PCT/US9511511
GENERAL INFORMATION:
APPLICANT: SEED, BRIAN
TITLE OF INVENTION: OVEREXPRESSION OF MAMMALIAN AND VIRAL
PROTEINS
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESS: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/11511
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: CLARK, PAUL T
REGISTRATION NUMBER: 40,162
REFERENCE/DOCKET NUMBER: 00786/226001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 1632 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
PCT-US95-11511-34

```

```

Query Match 3.5%; Score 31; DB 5; Length 1632;
Best Local Similarity 53.8%; Pred. No. 2.7;
Matches 64; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

```

```

OY 263 ctacacgcgcgtatgagtcgtctgctgagcctccttcgcaatccggtcttgc 322
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 555 ctccgcgcgcgtatgagtcgtctgctgagcctccttcgcaatccggtcttgc 496
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

```



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 4, 2002, 02:14:54 ; Search time 6206.58 Seconds  
(without alignments)  
867,408 Million cell updates/sec

Title: US-08-956-991-9\_COPY\_500\_1000

Perfect score: 501  
Sequence: 1 gcttcgcgcagcatcaata.....aacacagcagcttcagataac 501

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Capext 1.0

Searched: 11351937 seqs, 5372889281 residues

1 number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*  
1: em\_estfun:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estom:\*  
5: em\_estpl:\*  
6: em\_estba:\*  
7: em\_estro:\*  
8: em\_estov:\*  
9: em\_hcc:\*  
10: qb\_est1:\*  
11: qb\_est2:\*  
12: qb\_hcc:\*  
13: qb\_gss:\*  
14: em\_gss\_fun:\*  
15: em\_gss\_hum:\*  
16: em\_gss\_inv:\*  
17: em\_gss\_pln:\*  
18: em\_gss\_pro:\*  
19: em\_gss\_fod:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match Length	DB ID	Description
1	357.2	71.3	744 11	BF348620 602019924
2	261.2	52.1	528 11	BE721323 188457 MA
3	139.6	27.9	712 11	BE528580 602043580
4	86.4	17.2	1008 13	CNS020US
5	82.8	16.5	310 11	F13426
6	80.2	16.0	294 11	245894 HSC2X021 n
7	53	10.6	612 13	FR0023054
8	46.4	9.3	619 13	FR0030764
9	45.4	9.1	601 13	FR0022062
10	40	8.0	614 13	FR0030753
11	39.8	7.9	412 10	AI945592 bs14e12.y
12	39.8	7.9	617 11	BE976166 bs47f09.y

C 13	39.8	7.9	631 11	BE977242	BE977242 bs61d09.y
C 14	39.8	7.9	631 11	BF486312	BF486312 AT20576.5
C 15	39.8	7.9	646 11	BE976689	BE976689 bs54f11.y
C 16	39.8	7.9	822 11	BF486470	BF486470 AT20777.5
C 17	38.4	7.7	428 10	AI596938	AI596938 v148b08.y
C 18	37.8	7.5	359 10	AM481004	AM481004 34091 MAR
C 19	37.8	7.5	1097 13	CNS00815	AL051905 Drosoph11
C 20	37	7.4	608 11	BI082263	BI082263 602877444
C 21	36.8	7.3	249 11	R86463	R86463 RABEST0327
C 22	36.8	7.3	634 11	BI365325	BI365325 RES0593.5
C 23	36.8	7.3	635 11	BI237030	BI237030 RE33174.5
C 24	36.6	7.3	454 10	AI756515	AI756515 ELESSTea23
C 25	36.6	7.3	636 10	BE574841	BE574841 F03-P1 Tr
C 26	36.4	7.3	426 10	AA400616	AA400616 zu70C08.r
C 27	36.2	7.2	313 10	AI944477	AI944477 bs01h04.y
C 28	36.2	7.2	456 11	R31437	R31437 v170e10.r1
C 29	36.2	7.2	720 13	CNS023MG	AI179665 Tetradon
C 30	36	7.2	260 13	A2563606	A2563606 RPT-23-2
C 31	35.8	7.1	643 10	AA539574	AA539574 LD17768.5
C 32	35.8	7.1	648 11	BI214036	BI214036 RE19729.5
C 33	35.8	7.1	649 11	BI163320	BI163320 RE02690.5
C 34	35.8	7.1	655 11	BI167237	BI167237 RE07573.5
C 35	35.8	7.1	721 11	BI369358	BI369358 RES5515.5
C 36	35.8	7.1	1179 13	CNS040MR	AL269100 Tetradon
C 37	35.6	7.1	699 10	BE662781	BE662781 EST00533
C 38	35.6	7.1	704 10	BE662772	BE662772 EST00523
C 39	35.6	7.1	829 13	BH059023	BH059023 RPT-24-3
C 40	35.4	7.1	882 13	AF029541	AF029541 AF029541
C 41	35.4	7.1	1101 13	CNS016OW	AI107018 Drosoph11
C 42	35.2	7.0	425 10	AM208236	AM208236 M110889
C 43	35.2	7.0	475 10	AL381017	AI381017 MBR56A12
C 44	35.2	7.0	500 13	A2340922	A2340922 IM0073M05
C 45	35.2	7.0	535 11	BF004342	BF004342 EST432840

#### ALIGNMENTS

RESULT 1  
BF348620 744 bp mRNA EST 22-NOV-2000  
LOCUS 602019924F1 NCI-CCAP\_Brn67 Homo sapiens cDNA clone IMAGE:4155508  
DEFINITION 5', mRNA sequence.  
ACCESSION BF348620.1 GI:11296215  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 744)  
AUTHORS NIH-MGC htlp://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNLS Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: David N. Louis, M.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LAM9426 row: p column: 05  
High quality sequence stop: 697.  
Location/Qualifiers  
1..744  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4155508"  
/clone\_lib="NCI-CCAP\_Brn67"  
/tissue\_type="anaplastic oligodendroglioma with 1p/19q loss"

#### FEATURES

source  
1..744  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4155508"  
/clone\_lib="NCI-CCAP\_Brn67"  
/tissue\_type="anaplastic oligodendroglioma with 1p/19q loss"







ACCESSION	VERSION	KEYWORDS	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	COMMENT
C-2xa02, mRNA sequence.	F13426	F13426.1	GI:710043	EST.	Human.	Human.			
					Human sapiens				
					Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
					Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
					1 (bases 1 to 310)				
					Auffray, C., Behar, G., Bois, F., Bouchier, C., da Silva, C., Devignes				
					, M.D., Duprat, S., Houlgate, R., Juneau, M.N., Lamy, B., Lorenzo, F.,				
					Mitchell, H., Mariage-Samson, R., Pletu, G., Poullet, Y.,				
					Sebastien, H., and Tessier, A.				
					IMAGE: molecular integration of the analysis of the human genome				
					and its expression				
					C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)				
					95277534				
					Contact: Genethon				
					Genexpress-genethon				
					Genethon Centre de recherche sur le Genome Humain				
					1, rue de l'Internationale, Bp60 91002 I'RY Cedex, FRANCE				
					Tel: 33169472800				
					Fax: 33160778698				
					Email: genexpress@genethon.fr				
					Single read.				
					Genexpress_library_id: C; Genexpress_sequence_id: y2c-2xa02				
					Seq primer: (-21)M3-universal.				
					Location/Qualifiers				
					1..310				
					/organism="Homo sapiens"				
					/db_xref="taxon:9606"				
					/clone="C-2xa02"				
					/clone_lib="normalized infant brain cDNA"				
					/sex="Female"				
					/tissue_type="total brain"				
					/dev_stage="3 months old"				
					/note="Organ: brain; Vector: lambda BA; Site: 1: HindIII;				
					Site: 2: NotI; sex: Female; dev_stage=3 months old;				
					isolate-muscular atrophy patient; tissue_type=total brain"				
					: total mRNA was oligo-(dIy) primed and directionally				
					cloned 5' -> 3' into the HindIII -> NotI sites of the				
					lambda BA vector. Clone library from B.Souares, Psychiatry				
					Dept. Columbia University, USA.				
					Bento Soares, P.N.A.S in press"				
					Normalization_method:				
					79 a 73 c 89 g 66 t 3 others				
					BASE COUNT				
					ORIGIN				
					16.5%; Score 82.8; DB 11; Length 310;				
					87.1%; Pred. No. 9, 3e-13;				
					Matches 101; Conservative 0; Mismatches 14; Indels 1; Gaps 1.				
					16.5%; Score 82.8; DB 11; Length 310;				
					87.1%; Pred. No. 9, 3e-13;				
					Matches 101; Conservative 0; Mismatches 14; Indels 1; Gaps 1.				
					16.5%; Score 82.8; DB 11; Length 310;				
					87.1%; Pred. No. 9, 3e-13;				
					Matches 101; Conservative 0; Mismatches 14; Indels 1; Gaps 1.				
					16.5%; Score 82.8; DB 11; Length 310;				
					87.1%; Pred. No. 9, 3e-13;				
					Matches 101; Conservative 0; Mismatches 14; Indels 1; Gaps 1.				
					16.5%; Score 82.8; DB 11; Length 310;				
					87.1%; Pred. No. 9, 3e-13;				
					Matches 101; Conservative 0; Mismatches 14; Indels 1; Gaps 1.				
					16.5%; Score 82.8; DB 11; Length 310;				
					87.1%; Pred. No. 9, 3e-13;				
					Matches 101; Conservative 0; Mismatches 14; Indels 1; Gaps 1.				
					16.5%; Score 82.8; DB 11; Length 310;				
					87.1%; Pred. No. 9, 3e-13;				

REFERENCE	1 (Phases 1 to 294)
AUTHORS	Aufrey C., Behar G., Bois F., Bouchier C., da Silva C., Devignes M.D., Duprat S., Houigatte R., Juneau M.N., Lamy B., Lorenzo F., Mitchell H., Marije-Sanson R., Pietu G., Pouillot Y., Sebastian Kabakchis C. and Testier A.
TITLE	IMAGE: molecular integration of the analysis of the human genome and its expression
JOURNAL	C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)
MEDLINE	9527534
COMMENT	Contact: Genethon Genexpress-clenethon Genethon Centre de recherche sur le Genome Humain 1, rue de l'Internationale, Bldo 91002 Evry Cedex, FRANCE Tel : 33169472800 Fax : 33160778698 Email : genexpress@genethon.fr Single read. Genexpress.Library.Idt: C; Genexpress.sequence.Idt: ylc-zwh04 Seq primer: (-21)M1.universals Localization:Qualifiers
FEATURES	source 1..294 /organism "Homo sapiens" /db_xref="taxon:9606" /clone="C-zwh04" /clone_lib-"normalized infant brain cDNA" /sex="Female" /tissue_type-"total brain" /dev_stage-"3 months old" /note="Organ: brain; Vector: lambda BA; Site1: HindIII; Site2: NotI; sex=female; dev_stage=3 months old; isolate-muscular atrophy patient; tissue_type-total brain ; total mRNA was oligo-(dnt) primed and directionally cloned 5' -> 3' into the HindIII -> NotI sites of the lambda BA vector. Clone library from B.Souares, Psychiatry Dept. Columbia University, USA. Normalization_method: Benlo Soares, P.N.A.S in press"
BASE COUNT	79 a 66 c 87 g 58 t 4 others "
ORIGIN	
Query Match	16.0%; Score 80.2; DB 11; Length 294;
Best Local Similarity	89.7%; Pred. No. 4.9e-12;
Matches 96; Conservative	0; Mismatches 10; Indels 1; Gaps 1;
OY	395 ctgctggatcttggaagaaacgcaagcccgaggcttcagatcac 501       
Nb	2 CTGCTGGATCTTGAAGAAACCGCAAGCCGAGGCTTCGACATGCCAAG 60 
OY	455 agtttaactgaactgcctcagcagaataaacctggagcttcagatcac 501       
Db	61 AGTTTAAGCTGAATGCCTCATCAGCTAACACATACCCTGACTCAGTAC 107 
RESULT 7	
FR0022054	612 bp DNA GSS 10-DEC-1997
LOCUS	F.rubripes GSS sequence, clone 070M05aA12, genomic survey sequence
DEFINITION	AL014925
ACCESSION	AL014925.1 GI:2681293
VERSION	GSS: genome survey sequence.
KEYWORDS	Takifugu rubripes.
SOURCE	Takifugu rubripes.
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorphi; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Takifugu. 1 (bases 1 to 612) Eljar.G., Clark.M., Smith.S., Meek.S., Warner,S., Umranta.Y., Williams.G. and Brenner,S. Direct Submission Submitted (08-DEC-1997) MRC Human Genome Mapping Project Resource Centre Hinxton, Cambridge, CB10 1SB. Email: biohelp@hmp.mrc.ac.uk Comment: phnuescript IT KS



REFERENCE	ACANTHOMORPHA: Acanthopterygill; Percomorpha: Tetraodontiformes; Tetraodontidae: Takifugu.
1 (bases 1 to 614)	
Elgar, G., Clark, M., Smith, S., Meek, S., Warner, S., Umrانيا, Y., Williams, G., and Brenner, S.	
Direct Submission	
Submitted (09-JUN-1998) MRC Human Genome Mapping Project Resource Centre, Hinxton, Cambridge, CB10 1SB, UK. Email: biohelp@hmp.mrc.ac.uk	
Vector: pBluescript II KS	
VCtype: phagemid	
PRIMER: KS	
DESCR:	
One pass dye-terminator sequencing of cosmid cloned genomic sequence.	
FEATURES	
source	location/Qualifiers
	1..614
	/organism="Takifugu rubripes"
	/db_xref="taxon:31033"
	/clone_lib="cosmid 096F20"
	/clone="096F20cc11"
BASE COUNT	166 a 144 c 154 g 141 t 9 others
ORIGIN	
Query Match	8.0%; Score 40; DB 13; Length 614;
Best Local Similarity	63.7%; Pred. No. 1.1;
Matches	58; Conservative 0; Mismatches 33; Indels 0; Gaps 0;
Oy 1	gcttcgccagcatcattaccacccggagtgagtgatctgattgctggaatgacg 60
Db 443	gcttcgccagcatcattaccacccggagtgagtgatctgattgctggaatgacg 502
Oy 61	cggctgcccaatccatctcaatctgaa 91
Db 503	ngatgtccatcagcngnagtcacacgtcgca 533
RESULT 11	
A1945592/c	
LOCUS	bs14e12.y1 Drosophila melanogaster adult testis library Drosophila
DEFINITION	melanogaster cDNA clone bs14e12 5', mRNA sequence.
ACCESSION	A1945592
VERSION	A1945592.2 GI:9909325
KEYWORDS	EST.
SOURCE	Fruit fly.
ORGANISM	Drosophila melanogaster
REFERENCE	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
AUTHORS	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
	Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
	1 (bases 1 to 412)
	Andrews, J., Bouffard, G. G., Cheadle, C., Lu, J., Becker, K. G. and
	Oliver, B.
	Gene discovery using computational and microarray analysis of
	transcription in the drosophila melanogaster testis
	Genome Res. 10 (12), 2030-2043 (2000)
	On Aug 17, 1999 this sequence version replaced gi:5735990.
	Contact: Brian Oliver
	Laboratory of Cellular and Developmental Biology
	NIDDK, National Institutes of Health
	6 Center Drive MSC 2715, Bldg 6, Rm B1-13, Bethesda, MD 20892 USA
	Fax: (301) 496 5239
	Email: oliver@helix.nih.gov,
	http://www.niddd.nih.gov/intlam/people/people/boliver.htm
	Tissue isolation and library construction performed at the National
	Institute of Diabetes and Digestive and Kidney Diseases, NIH (see
	http://www.niddd.nih.gov/intlam/people/boliver.htm). DNA sequencing
	and analyses performed by National Institutes of Health Intramural
	Sequencing Center (NISC; see http://www.nisc.nih.gov).
	Seq primer: M13RP1 reverse primer (AB1).

FEATURES	source	location/Qualifiers
1..412		
/organism="Drosophila melanogaster"		
/strain="y1" w167c11/y"		
/db_xref="taxon:7227"		
/clone="bs14e12"		
/clone.lib="Drosophila melanogaster adult testis library"		
/sex="male"		
/dev_stage="1-5 day adult"		
/lab_host="SOLR (Stratagene)"		
/note="Organ: testis; Vector: pBluescript SK (Stratagene); Site-1: EcoR 1; Site-2: Xho 1; Testes dissected from 1-5 day adult y1" w167c11/y males raised at 25°C. RNA isolated using Trizol (Life Technologies) and a single round of poly(A)+ selection using Stratagene ZAP-CDNA synthesis kit. Oligo dt-primed, size fractionated ~1-6 kb, and directionally cloned at EcoRI and XhoI in Uni-ZAP XR. Following a single round of amplification pBluescript SK plasmids were mass excised. A distribution channel for clones is being sought, but not currently available. Requests for clones cannot be honored."		
BASE COUNT	114 a	120 c 96 g 82 t
ORIGIN		
Query Match	7.9%; Score 39.8; DB 10; Length 412;	
Best Local Similarity	62.6%; Pred. No. 1.1;	
Matches	62; Conservative	0; Mismatches 37; Indels 0; Gaps 0.
Oy	345	tcgtgaccatctcgcgacatctgtgctcgggtctactgctcttgcgtcgcgtg 404
Db	187	TGCTGTCCTCGCTCGCTCGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 128
Oy	405	tgcgagagagc 443
Db	127	TGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 89
RESULT 12		
BE976166/c	617 bp	EST 04-OCT-2000
LOCUS	bs47f09.y1 Drosophila melanogaster adult testis library Drosophila	
DEFINITION	melanogaster cDNA clone bs47f09 5', mRNA sequence.	
ACCESSION	BE976166	
VERSION	BE976166.1	GI:10605489
KEYWORDS	EST.	
SOURCE	fruit fly.	
ORGANISM	Drosophila melanogaster	
	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;	
	Muscomorpha; Ephemeroptera; Diptera; Brachycera;	
	1 (bases 1 to 617)	
REFERENCE	Andrews, J., Boultard, G. and Oliver, B.	
AUTHORS	Drosophila melanogaster testis expressed sequence tags	
TITLE	Unpublished (1999)	
JOURNAL	Contact: Brian Oliver	
COMMENT	Laboratory of Cellular and Developmental Biology	
	NIDDK, National Institutes of Health	
	6 Center Drive MSC 2715, Bldg 6, Rm B1-13, Bethesda, MD 20892 USA	
	Fax: (301) 496 5239	
	Email: oliverb@helix.nih.gov,	
	http://www.nidk.nih.gov/intram/people/bolliver.htm	
	Tissue isolation and library construction performed at the National	
	Institute of Diabetes and Digestive and Kidney Diseases, NIH (see	
	http://www.nidk.nih.gov/intram/people/bolliver.htm). DNA sequencing	
	and analyses performed by National Institutes of Health Intramural	
	Sequencing Center (NISC; see http://www.nisc.nih.gov).	
	Plate: 47 row: f column: 09	
	Seq primer: M13p1 reverse primer (ABI).	
FEATURES	Location/Qualifiers	
source	1..617	
	/organism="Drosophila melanogaster"	



BASE COUNT  
ORIGIN

157 a 202 c 167 g 105 t

/sex="male"  
/dev stage="0-3 day old Ore-R males"  
/lab host="Plates AT\_10-AT\_120: DH5-alpha. plates  
AT\_121-AT\_319: DH5-alpha."juva"  
/note="Organ: ADULT testes; Vector: pOTB7; Site: 1; Exon:1  
Site\_2: Xho1; The mRNA for the testis library was made  
from testes and seminal vesicles hand dissected from 0-3  
day old Ore-R males. RNA kindly provided by the lab of  
Margaret Fuller. Sized fractionated cDNAs were directly  
ligated into pOTB7. Plasmid cDNA library."

Query Match:	7.9%	Score 39.8:	DB 11:	Length 631:
Best Local Similarity:	62.6%:	Pred. No. 1.3:		
Matches 62:	Conservative	0:	Mismatches 37:	Indels 0:
			Gaps	0:
OY	345	tcgtgacacatctctcgcacatcgtgctcgtgaggtctctactctgcttcttctgctcgtgcttg	404	
	513	tgtctgtctccctgcgtcgcgcttgcctatgagtcctcctcctcgtctgtgctgctgctcctgc	454	
OY	405	tcctgaggaacatctgacagagacagagctgaaagagcttga	443	
Db	453	tgtctgtgtgactctgcctcgtctgactctcctcattaacctga	415	

RESULT	15
LOCUS	BE976689/c
DEFINITION	BE976689 646 bp mRNA EST 04-OCT-2000 bs54f11.y1 Drosophila melanogaster adult testis library Drosophilla
ACCESSION	melanogaster cDNA clone bs54f11 5', mRNA sequence. BE976689
VERSION	BE976689.1 GI:10606415
KEYWORDS	EST.
SOURCE	Fruit fly.
ORGANISM	Drosophila melanogaster Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila. 1 (bases 1 to 646) Andrews, J., Bouffard, G., and Oliver, B. Drosophila melanogaster testis expressed sequence tags unpublished (1999)
REFERENCE	Contact: Brian Oliver
AUTHORS	
TITLE	
JOURNAL	
COMMENT	

FEATURES  
SOURCE

Private MS22715, Bldg 6, Rm B1-15, Bethesda, MD 20892 USA  
 496.5239  
 ivermectin.nih.gov,  
 www.niddk.nih.gov/intram/people/poliver.htm  
 isolation and library construction performed at the National  
 of Diabetes and Digestive and Kidney Diseases, NIH (see  
 www.niddk.nih.gov/intram/people/poliver.htm). DNA sequencing  
 uses performed by National Institutes of Health Intramural  
 Center (NISC; see <http://www.nisc.nih.gov>).  
 4 row: 1 column: 11  
 Location/Qualifiers  
 1. 646  
 /organism="Protophila melanogaster"  
 /strain="y[\*] w[67c1] y"  
 /db\_xref="taxon:7227"  
 /clone="hs54f11"  
 /clone\_11b="Protophila melanogaster adult testis library"  
 /sex="male"  
 /dev\_stage="1-5 day adult"  
 /lab\_host="Solk (Stratagene)"  
 /note="Organ testis; Vector: plunscript SK (Stratagene)"  
 Site: 1: Ecor 1; Site: 2: Xho 1; Testis dissected from 1-5  
 day adult y[\*] w[67c1] y males raised at 25°C. RNA  
 isolated using Trizol (Life Technologies) and a single  
 round of Poly(A)<sup>+</sup> selection using Oligotex (Qiagen). cDNA

BASE COUNT  
ORIGIN

184 a 143 c 153 g 166 t

[illegible]

```
Search completed: March  4, 2002, 02:15:04
Job time: 6510 sec
```

gene ZAP-CDNA synthesis  
located -1.6 kb, and  
d XhoI in Uni-ZAP XR.  
Identification plasmid SK  
distribution channel for  
currently available.

```
Length 646;
Index      0; Gaps      0.
ttgtgcttcgcagtg 404
|||||
ggcgcgcgcgcgcgcgc 382
```





[illegible]

XX This polynucleotide comprises cDNA fragment E51 that was isolated  
CC from a human trisomy 21 foetal brain (14-wk gestation) cDNA  
CC library following a modified direct cDNA selection technique  
CC applied to bacterial and pl artificial chromosomes between ET52  
CC and MX1. E51 was used as a probe to screen the trisomy 21 foetal  
CC brain library. 62 clones were isolated from the 2 million clones  
CC of the original library. overlapping clones were obtained that  
CC were used to deduce a full-length sequence (see AAV31981) coding  
CC for novel bowl syndrome-rev1 adhesion molecule DS-CAM1 (see  
CC AAV42086). A splice variant, DS-CAM2 (see AAV31988), was also  
CC identified. DS-CAM polypeptides are associated with developmental  
CC and neurological processes. The polypeptides and nucleic acids  
CC are used to develop products for the detection, diagnosis and  
CC therapy of developmental and neurological abnormalities.  
XX  
XX Sequence 388 BP; 83 A; 119 C; 95 G; 91 T; 0 other;  
SQ

Query Match:	60.7%	Score 304.02	DB 19	Length 388	
Best Local Similarity	89.5%	Pred. No. 3-2e-86			
Matches 350	Conservative	0	Mismatches 38	Indels 3	Gaps 2
QY 100	ctttgagctatcagctgagctacagctcagctgagactcccttccaaagtccaaacalc	159			
DB 388	ctttggactacactgtttggacacacagctcagagagactctctctccaaagcct-acatcc	330			
QY 160	tgatgactctcgaagaagctcagctgatacgaactgcagatgagatgagatgctgcaacagccg	219			
DB 329	tgcatgacctgcagaggaagcctcctgtatgagctccagatgcggctgcacacagtcggg	270			
QY 220	gcttgctgagatgaatcgaatcgaatcttgcacagctgaactlaagatggcaatcaatccctc	279			
DB 269	gcctgcgcgcagagagcagcctcactctctcagcctgaactcagatgcagctacatctctc	210			
QY 280	cactcattcaatcagctgctcctcaadaagcaagaagctgacacaaccaagagagctcaa	339			
DB 209	cactcatttaagctcagcttgcctc-caaagcgaagagcgctcagcagcaacagagcgctcaa	152			
QY 340	gactcctcgtacacatctcctcctatcctgctgagctgctctacgctcttctgctctgct	399			
DB 151	gagctgctgagcactctctctgtatctctgctgagctgctgctctgctgctgctgctgct	92			
QY 400	gcttgctgatt	459			
DB 91	cgcttgctgatt	32			
QY 460	agctgag	490			
DB 31	agctgag	1			
RESULT 4					
AAV31988					
ID AAV31988	standard: cDNA; 6413 BP.				
AC AAV31988:					
XX 28-SEP-1998	(first entry)				
DE Human Down syndrome-cell adhesion molecule DS-CAM2 cDNA.					
XX DS-CAM2: Down syndrome-cell adhesion molecule; neural cells;					
KW signal transduction; trisomy 21; mental retardation;					
KW holoprosencephaly; corpus callosum agenesis;					
KW schizencephaly; diagnosis; human; ds; ss.					
XX Homo sapiens.					
XX Key location/qualifiers					
PH 453...5168					
FT CDS					
XX /*lag- a					







```

PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 28-JUL-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 13-SEP-1999; 99US-0154018.
PR 15-SEP-1999; 99US-0154039.
PR 16-SEP-1999; 99US-0154779.
PR 20-SEP-1999; 99US-0155139.
PR 22-SEP-1999; 99US-0155486.
PR 23-SEP-1999; 99US-0155659.
PR 24-SEP-1999; 99US-0156458.
PR 28-SEP-1999; 99US-0156596.
PR 29-SEP-1999; 99US-0157117.
PR 04-OCT-1999; 99US-0157753.
PR 05-OCT-1999; 99US-0157865.
PR 06-OCT-1999; 99US-0158029.
PR 07-OCT-1999; 99US-0158232.
PR 08-OCT-1999; 99US-0158369.
PR 12-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.

```

```

PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

```

Query Match 7 1%; Score 45.6; DB 21; Length 1627;
Best Local Similarity 52.7%; Pred. No. 0.41; Mismatches 69; Indels 0; Gaps 0;
Matches 77; Conservative 0;

```

```

QY 338 aagatccctgacacatcc'gcatcctgctggggtctactgctcttctctg 397
DB 471 AACGCTTGACATCATCTCCAGATCCAGCTTGCTGCTGCTGCTG 412
QY 398 ctggttgcgagagagagagagagagagagagagagagagagag 457
DB 411 CTGCTTGATAGACCCGCTATTCTGAGAGAGATCATCTCTCTAGACGACCTTGATT 352
QY 458 ttagctgaatgctcagagagagagagagagagagagagagagag 483
DB 351 ATACGAGACATTTGATGAGAGATACA 326

```

```

RESULT 7
AAC1871/c
ID AAC1871 standard; DNA: 1679 bp.
XX AC AAC1871:
XX 17-OCT-2000 (first entry)
XX DT
XX DE Arabidopsis thaliana DNA fragment SK011111.
XX XX
XX KW Hybridisation assay; gene expression control;
XX KW protein identification; signal transduction pathway;
XX PN metabolic pathway; promoter; termination sequence; ss.
XX Arabidopsis thaliana.
XX EP1033405-AZ.
PD 06-SEP-2000.
XX XX
XX 25-FEB-2000; 2000EP-0301419.
XX PF
XX 25-FEB-1999; 99US-0121825.
XX PR 05-MAR-1999; 99US-0123180.
XX PR 09-MAR-1999; 99US-0123548.
XX PR 23-MAR-1999; 99US-0125788.
XX PR 25-MAR-1999; 99US-0126264.
XX PR 29-MAR-1999; 99US-0126785.
XX PR 01-APR-1999; 99US-0127462.
XX PR 06-APR-1999; 99US-0128234.
XX PR 08-APR-1999; 99US-0128714.
XX PR 16-APR-1999; 99US-0129845.
XX PR 19-APR-1999; 99US-0130077.
XX PR 21-APR-1999; 99US-0130449.

```

PR 23-APR-1999; 99US-0130510.  
 PR 23-APR-1999; 99US-0130891.  
 PR 28-APR-1999; 99US-0131449.  
 PR 30-APR-1999; 99US-0132048.  
 PR 30-APR-1999; 99US-0132407.  
 PK 04-MAY-1999; 99US-0132484.  
 PK 05-MAY-1999; 99US-0132485.  
 PK 06-MAY-1999; 99US-0132486.  
 PR 06-MAY-1999; 99US-0132487.  
 PR 07-MAY-1999; 99US-0132863.  
 PR 11-MAY-1999; 99US-0134256.  
 PR 14-MAY-1999; 99US-0134218.  
 PR 14-MAY-1999; 99US-0134221.  
 PR 14-MAY-1999; 99US-0134370.  
 PR 18-MAY-1999; 99US-0134378.  
 PR 19-MAY-1999; 99US-0134941.  
 PR 20-MAY-1999; 99US-0135124.  
 PR 24-MAY-1999; 99US-0135353.  
 PR 25-MAY-1999; 99US-0135629.  
 PR 27-MAY-1999; 99US-0136021.  
 PR 28-MAY-1999; 99US-0136382.  
 PR 01-JUN-1999; 99US-0137222.  
 PR 03-JUN-1999; 99US-0137528.  
 PR 04-JUN-1999; 99US-0137702.  
 PR 07-JUN-1999; 99US-0137724.  
 PR 08-JUN-1999; 99US-0138099.  
 PR 10-JUN-1999; 99US-0138540.  
 PR 10-JUN-1999; 99US-0138847.  
 PR 14-JUN-1999; 99US-0139119.  
 PR 16-JUN-1999; 99US-0139452.  
 PR 16-JUN-1999; 99US-0139453.  
 PR 17-JUN-1999; 99US-0139492.  
 PR 18-JUN-1999; 99US-0139454.  
 PR 18-JUN-1999; 99US-0139455.  
 PR 18-JUN-1999; 99US-0139456.  
 PR 18-JUN-1999; 99US-0139457.  
 PR 18-JUN-1999; 99US-0139458.  
 PR 18-JUN-1999; 99US-0139459.  
 PR 18-JUN-1999; 99US-0139460.  
 PR 18-JUN-1999; 99US-0139461.  
 PR 18-JUN-1999; 99US-0139462.  
 PR 18-JUN-1999; 99US-0139463.  
 PR 18-JUN-1999; 99US-0139750.  
 PR 18-JUN-1999; 99US-0139763.  
 PR 21-JUN-1999; 99US-0139817.  
 PR 22-JUN-1999; 99US-0139899.  
 PR 23-JUN-1999; 99US-0140353.  
 PR 23-JUN-1999; 99US-0140354.  
 PR 24-JUN-1999; 99US-0140695.  
 PR 28-JUN-1999; 99US-0140823.  
 PR 29-JUN-1999; 99US-0140991.  
 PR 30-JUN-1999; 99US-0141287.  
 PR 01-JUL-1999; 99US-0141842.  
 PR 01-JUL-1999; 99US-0142154.  
 PR 02-JUL-1999; 99US-0142055.  
 PR 06-JUL-1999; 99US-0142390.  
 PR 08-JUL-1999; 99US-0142803.  
 PR 09-JUL-1999; 99US-0142920.  
 PR 12-JUL-1999; 99US-0142977.  
 PR 13-JUL-1999; 99US-0143542.  
 PR 14-JUL-1999; 99US-0143624.  
 PR 15-JUL-1999; 99US-0144005.  
 PR 16-JUL-1999; 99US-0144085.  
 PR 16-JUL-1999; 99US-0144086.  
 PR 19-JUL-1999; 99US-0144325.  
 PR 19-JUL-1999; 99US-0144331.  
 PR 19-JUL-1999; 99US-0144332.  
 PR 19-JUL-1999; 99US-0144333.  
 PR 19-JUL-1999; 99US-0144334.  
 PR 19-JUL-1999; 99US-0144335.  
 PR 20-JUL-1999; 99US-0144352.

PR 20-JUL-1999; 99US-0144632.  
 PR 20-JUL-1999; 99US-0144884.  
 PR 21-JUL-1999; 99US-0144814.  
 PR 21-JUL-1999; 99US-0145086.  
 PR 21-JUL-1999; 99US-0145088.  
 PR 22-JUL-1999; 99US-0145085.  
 PR 22-JUL-1999; 99US-0145087.  
 PR 22-JUL-1999; 99US-0145089.  
 PR 22-JUL-1999; 99US-0145192.  
 PR 23-JUL-1999; 99US-0145145.  
 PR 23-JUL-1999; 99US-0145218.  
 PR 23-JUL-1999; 99US-0145224.  
 PR 26-JUL-1999; 99US-0145276.  
 PR 27-JUL-1999; 99US-0145913.  
 PR 27-JUL-1999; 99US-0145918.  
 PR 27-JUL-1999; 99US-0145919.  
 PR 28-JUL-1999; 99US-0145951.  
 PR 02-AUG-1999; 99US-0146386.  
 PR 02-AUG-1999; 99US-0146389.  
 PR 02-AUG-1999; 99US-0146389.  
 PR 03-AUG-1999; 99US-0147038.  
 PR 04-AUG-1999; 99US-0147204.  
 PR 04-AUG-1999; 99US-0147302.  
 PR 05-AUG-1999; 99US-0147192.  
 PR 05-AUG-1999; 99US-0147260.  
 PR 06-AUG-1999; 99US-0147303.  
 PR 06-AUG-1999; 99US-0147416.  
 PR 09-AUG-1999; 99US-0147493.  
 PR 09-AUG-1999; 99US-0147935.  
 PR 10-AUG-1999; 99US-0148171.  
 PR 11-AUG-1999; 99US-0148319.  
 PR 12-AUG-1999; 99US-0148341.  
 PR 13-AUG-1999; 99US-0148565.  
 PR 13-AUG-1999; 99US-0148684.  
 PR 16-AUG-1999; 99US-0149368.  
 PR 17-AUG-1999; 99US-0149175.  
 PR 18-AUG-1999; 99US-0149426.  
 PR 20-AUG-1999; 99US-0149722.  
 PR 20-AUG-1999; 99US-0149723.  
 PR 20-AUG-1999; 99US-0149902.  
 PR 23-AUG-1999; 99US-0149930.  
 PR 25-AUG-1999; 99US-0150566.  
 PR 26-AUG-1999; 99US-0150884.  
 PR 27-AUG-1999; 99US-0151065.  
 PR 27-AUG-1999; 99US-0151086.  
 PR 30-AUG-1999; 99US-0151303.  
 PR 31-AUG-1999; 99US-0151438.  
 PR 01-SEP-1999; 99US-0151940.  
 PR 07-SEP-1999; 99US-0152363.  
 PR 10-SEP-1999; 99US-0153070.  
 PR 13-SEP-1999; 99US-0153758.  
 PR 15-SEP-1999; 99US-0154018.  
 PR 16-SEP-1999; 99US-0154039.  
 PR 20-SEP-1999; 99US-0154779.  
 PR 22-SEP-1999; 99US-0155139.  
 PR 23-SEP-1999; 99US-0155486.  
 PR 24-SEP-1999; 99US-0155659.  
 PR 28-SEP-1999; 99US-0156458.  
 PR 29-SEP-1999; 99US-0156596.  
 PR 04-OCT-1999; 99US-0157117.  
 PR 05-OCT-1999; 99US-0157753.  
 PR 06-OCT-1999; 99US-0157865.  
 PR 07-OCT-1999; 99US-0158029.  
 PR 08-OCT-1999; 99US-0158232.  
 PR 12-OCT-1999; 99US-0158369.  
 PR 13-OCT-1999; 99US-0158293.  
 PR 13-OCT-1999; 99US-0158294.  
 PR 13-OCT-1999; 99US-0159295.  
 PR 14-OCT-1999; 99US-0159329.  
 PR 14-OCT-1999; 99US-0159330.  
 PR 14-OCT-1999; 99US-0159331.

[illegible]

DR	P-PSDB; AAKR3943.
XX	Rabbit skeletal muscle calcium channel gamma-subunit encoding
PR	cDNA - useful for identification of agonists-antagonists, design
PT	of therapeutic agents and diagnostic assays for e.g. Lambert-Eaton syndrome, etc.
XX	
PS	Disclosure; Fig 1: 104pp; English.
XX	
CC	cDNA clones encoding each of the four CC subunits (alpha), alpha2,
CC	beta, gamma) are now available. Various combinations of such cDNAs
CC	may be incorporated into recombinant expression systems. This
CC	allows assessment of the contribution of the different subunits to
CC	the transport and regulation of calcium ions. Availability of the
CC	cDNA and its prod. is useful in the design of therapeutic agents and
CC	diagnostic assays for diseases such as Lambert-Eaton Syndrome, involving autoimmune interactions with CCS.
XX	
SQ	Sequence 1171 BP; 202 A; 407 C; 132 G; 230 T; 0 other;
Query Match	6, 7%; Score 33.8; DB 12; Length 1171;
Best Local Similarity	60.2%; Pred No. 1.3;
Matches	56; Conservative 0; Mismatches 37; Indels 0; Gaps 0;
Oy	308 gaagaaaggcgcgaacacaaagaagggctcaagatcccgtagaccattctgatactg 367   Db 40 gagcgcgcacgatcccgcgaagaagccccaaaagtcggtagacctctctgatactg 99  Oy 368 gtcgggttctactatgcccttttgagcttcagt 400   Db 100 gtgcgcatccgltgclygeratvaagggcggg 132
RESULT	9
ID	AAC53185 standard; DNA: 692 BP.
XX	AAC53185:
XX	
DT	18-OCT-2000 (first entry)
XX	
DE	Arabidopsis thaliana DNA fragment SKU ID NO: 73663.
KX	Hybridisation assay; gene for mapping; gene expression control;
KW	protein identification; signal transduction pathway;
KM	metabolic pathway; promoter; translation sequence; ss.
XX	
OS	Arabidopsis thaliana.
XX	
FN	EPI033405-AZ.
XX	
PD	06-SEP-2000.
XX	
PF	25-FEB-2000; 2000EP-0301439.
XX	
PR	25-FEB-1999; 99US-0121825.
PR	05-MAR-1999; 99US-0123180.
PR	09-MAR-1999; 99US-0123548.
PR	23-MAR-1999; 99US-0125788.
PR	25-MAR-1999; 99US-0126264.
PR	29-MAR-1999; 99US-0126785.
PR	01-APR-1999; 99US-0127462.
PR	06-APR-1999; 99US-0128234.
PR	08-APR-1999; 99US-0128714.
PR	16-APR-1999; 99US-0129845.
PR	19-APR-1999; 99US-0130077.
PR	21-APR-1999; 99US-0130449.
PR	23-APR-1999; 99US-0130510.
PR	23-APR-1999; 99US-0130891.
PR	28-APR-1999; 99US-0131449.
PR	30-APR-1999; 99US-0132048.
PR	30-APR-1999; 99US-0132407.

PR 04-MAY-1999; 9905-0132484.  
PR 05-MAY-1999; 9905-0132485.  
PR 06-MAY-1999; 9905-0132486.  
PR 07-MAY-1999; 9905-0132487.  
PR 11-MAY-1999; 9905-0132488.  
PR 14-MAY-1999; 9905-0134219.  
PR 14-MAY-1999; 9905-0134221.  
PR 14-MAY-1999; 9905-0134370.  
PR 18-MAY-1999; 9905-0134768.  
PR 19-MAY-1999; 9905-0134941.  
PR 20-MAY-1999; 9905-0135124.  
PR 21-MAY-1999; 9905-0135353.  
PR 24-MAY-1999; 9905-0135629.  
PR 25-MAY-1999; 9905-0136021.  
PR 27-MAY-1999; 9905-0136392.  
PR 28-MAY-1999; 9905-0136782.  
PR 01-JUN-1999; 9905-0137222.  
PR 03-JUN-1999; 9905-0137528.  
PR 04-JUN-1999; 9905-0137502.  
PR 07-JUN-1999; 9905-0137724.  
PR 08-JUN-1999; 9905-0138094.  
PR 10-JUN-1999; 9905-0138540.  
PR 10-JUN-1999; 9905-0138847.  
PR 14-JUN-1999; 9905-0139119.  
PR 16-JUN-1999; 9905-0139452.  
PR 16-JUN-1999; 9905-0139453.  
PR 17-JUN-1999; 9905-0139454.  
PR 18-JUN-1999; 9905-0139455.  
PR 18-JUN-1999; 9905-0139456.  
PR 18-JUN-1999; 9905-0139457.  
PR 18-JUN-1999; 9905-0139458.  
PR 18-JUN-1999; 9905-0139459.  
PR 18-JUN-1999; 9905-0139460.  
PR 18-JUN-1999; 9905-0139461.  
PR 18-JUN-1999; 9905-0139462.  
PR 18-JUN-1999; 9905-0139463.  
PR 18-JUN-1999; 9905-0139750.  
PR 18-JUN-1999; 9905-0139763.  
PR 21-JUN-1999; 9905-0139817.  
PR 22-JUN-1999; 9905-0139899.  
PR 22-JUN-1999; 9905-0140353.  
PR 23-JUN-1999; 9905-0140354.  
PR 24-JUN-1999; 9905-0140695.  
PR 28-JUN-1999; 9905-0140823.  
PR 29-JUN-1999; 9905-0140991.  
PR 30-JUN-1999; 9905-0141287.  
PR 01-JUL-1999; 9905-0141842.  
PR 02-JUL-1999; 9905-0142154.  
PR 06-JUL-1999; 9905-0142055.  
PR 08-JUL-1999; 9905-0142380.  
PR 09-JUL-1999; 9905-0142803.  
PR 12-JUL-1999; 9905-0142920.  
PR 13-JUL-1999; 9905-0142977.  
PR 14-JUL-1999; 9905-0143642.  
PR 15-JUL-1999; 9905-0143644.  
PR 15-JUL-1999; 9905-0144005.  
PR 16-JUL-1999; 9905-0144085.  
PR 16-JUL-1999; 9905-0144086.  
PR 19-JUL-1999; 9905-0144325.  
PR 19-JUL-1999; 9905-0144331.  
PR 19-JUL-1999; 9905-0144332.  
PR 19-JUL-1999; 9905-0144333.  
PR 19-JUL-1999; 9905-0144334.  
PR 19-JUL-1999; 9905-0144335.  
PR 20-JUL-1999; 9905-0144352.  
PR 20-JUL-1999; 9905-0144632.  
PR 20-JUL-1999; 9905-0144884.  
PR 21-JUL-1999; 9905-0144881.  
PR 21-JUL-1999; 9905-0145086.  
PR 21-JUL-1999; 9905-0145088.  
  
PR 22-JUL-1999; 9905-0145085.  
PR 22-JUL-1999; 9905-0145087.  
PR 22-JUL-1999; 9905-0145089.  
PR 23-JUL-1999; 9905-0145192.  
PR 23-JUL-1999; 9905-0145145.  
PR 23-JUL-1999; 9905-0145218.  
PR 23-JUL-1999; 9905-0145224.  
PR 26-JUL-1999; 9905-0145276.  
PR 27-JUL-1999; 9905-0145913.  
PR 27-JUL-1999; 9905-0145918.  
PR 28-JUL-1999; 9905-0145951.  
PR 02-AUG-1999; 9905-0146386.  
PR 02-AUG-1999; 9905-0146388.  
PR 02-AUG-1999; 9905-0146389.  
PR 03-AUG-1999; 9905-0147048.  
PR 04-AUG-1999; 9905-0147204.  
PR 04-AUG-1999; 9905-0147302.  
PR 05-AUG-1999; 9905-0147192.  
PR 05-AUG-1999; 9905-0147260.  
PR 06-AUG-1999; 9905-0147303.  
PR 06-AUG-1999; 9905-0147416.  
PR 09-AUG-1999; 9905-0147493.  
PR 09-AUG-1999; 9905-0147945.  
PR 10-AUG-1999; 9905-0148171.  
PR 11-AUG-1999; 9905-0148319.  
PR 12-AUG-1999; 9905-0148341.  
PR 13-AUG-1999; 9905-0148565.  
PR 13-AUG-1999; 9905-0148684.  
PR 16-AUG-1999; 9905-0149368.  
PR 17-AUG-1999; 9905-0149175.  
PR 18-AUG-1999; 9905-0149426.  
PR 20-AUG-1999; 9905-0149722.  
PR 20-AUG-1999; 9905-0149723.  
PR 20-AUG-1999; 9905-0149929.  
PR 20-AUG-1999; 9905-0149902.  
PR 23-AUG-1999; 9905-0149930.  
PR 25-AUG-1999; 9905-0150566.  
PR 26-AUG-1999; 9905-0150884.  
PR 27-AUG-1999; 9905-0151065.  
PR 27-AUG-1999; 9905-0151066.  
PR 27-AUG-1999; 9905-0151080.  
PR 30-AUG-1999; 9905-0151303.  
PR 31-AUG-1999; 9905-0151438.  
PR 01-SEP-1999; 9905-0151910.  
PR 07-SEP-1999; 9905-0152463.  
PR 10-SEP-1999; 9905-0154070.  
PR 13-SEP-1999; 9905-0154078.  
PR 15-SEP-1999; 9905-0154079.  
PR 16-SEP-1999; 9905-0154079.  
PR 20-SEP-1999; 9905-0155149.  
PR 22-SEP-1999; 9905-0155486.  
PR 23-SEP-1999; 9905-0155659.  
PR 24-SEP-1999; 9905-0155659.  
PR 28-SEP-1999; 9905-0156458.  
PR 29-SEP-1999; 9905-0156596.  
PR 04-OCT-1999; 9905-0157117.  
PR 05-OCT-1999; 9905-0157753.  
PR 06-OCT-1999; 9905-0157865.  
PR 07-OCT-1999; 9905-0158029.  
PR 08-OCT-1999; 9905-0158232.  
PR 12-OCT-1999; 9905-0158369.  
PR 13-OCT-1999; 9905-0159293.  
PR 13-OCT-1999; 9905-0159294.  
PR 13-OCT-1999; 9905-0159295.  
PR 14-OCT-1999; 9905-0159329.  
PR 14-OCT-1999; 9905-0159330.  
PR 14-OCT-1999; 9905-0159331.  
PR 14-OCT-1999; 9905-0159637.  
PR 14-OCT-1999; 9905-0159638.  
PR 18-OCT-1999; 9905-0159584.  
PR 21-OCT-1999; 9905-0160741.  
PR 21-OCT-1999; 9905-0160767.

```
PR 21-OCT-1999: 99US-0160768.
PR 21-OCT-1999: 99US-0160770.
PR 21-OCT-1999: 99US-0160814.
PR 21-OCT-1999: 99US-0160815.
PR 22-OCT-1999: 99US-0160980.
PR 22-OCT-1999: 99US-0160981.
PR 22-OCT-1999: 99US-0160989.
PR 25-OCT-1999: 99US-0161404.
PR 25-OCT-1999: 99US-0161405.
PR 25-OCT-1999: 99US-0161406.
PR 26-OCT-1999: 99US-0161359.
PR 26-OCT-1999: 99US-0161360.
PR 26-OCT-1999: 99US-0161361.
PR 28-OCT-1999: 99US-0161920.
PR 28-OCT-1999: 99US-0161922.
PR 28-OCT-1999: 99US-0161993.
PR 28-OCT-1999: 99US-0161993.
PR 29-OCT-1999: 99US-0162142.
```

```
Query Match 6.6%; Score 33; DB 21: Length 692;
          1st Local Similarity 60.7%; Pred. No. 1.8;
          .ches 54; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
```

```
OY 373 ggtctactgtctcttctgtctgtctgtgtgagagagagagagagagct 432
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 347 gggctaatgactcaatcaccgatacgcgactgtctcgaggaactgcccagcagcttga 406
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 433 gaagagctgagagatgcaagaagcttag 461
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 407 gtggttgataagaagctcaagcattcag 435
```

```
RESULT 10
AAC38550
ID AAC38550 standard: DNA: 533 Bp.
```

XX AAC38550:

XX 17-OCT-2000 (first entry)

XX Arabidopsis thaliana DNA fragment SKO ID No: 21378.

XX Hybridisation assay: genetic mapping; gene expression control;

KM protein identification: signal transduction pathway;

KW metabolic pathway; promoter; termination sequence; ss.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000: 2000EP-0301439.

XX 25-FEB-1999: 99US-0121825.

XX 09-MAR-1999: 99US-0123180.

XX 23-MAR-1999: 99US-0123548.

XX 23-MAR-1999: 99US-0125788.

XX 23-MAR-1999: 99US-0126264.

XX 29-MAR-1999: 99US-0126785.

XX 01-APR-1999: 99US-0127462.

XX 06-APR-1999: 99US-0128234.

XX 08-APR-1999: 99US-0128714.

XX 16-APR-1999: 99US-0129845.

XX 19-APR-1999: 99US-0130077.

XX 21-APR-1999: 99US-0130449.

XX 23-APR-1999: 99US-0130510.

XX 23-APR-1999: 99US-0130891.

XX 28-APR-1999: 99US-0131449.

XX 30-APR-1999: 99US-0132048.

XX 30-APR-1999: 99US-0132407.

XX 04-MAY-1999: 99US-0132484.

XX 05-MAY-1999: 99US-0132485.

XX 06-MAY-1999: 99US-0132486.

```
PR 06-MAY-1999: 99US-0132487.
PR 07-MAY-1999: 99US-0132863.
PR 11-MAY-1999: 99US-0134256.
PR 14-MAY-1999: 99US-0134218.
PR 14-MAY-1999: 99US-0134219.
PR 14-MAY-1999: 99US-0134270.
PR 14-MAY-1999: 99US-0134370.
PR 18-MAY-1999: 99US-0134768.
PR 19-MAY-1999: 99US-0134941.
PR 20-MAY-1999: 99US-0135124.
PR 21-MAY-1999: 99US-0135353.
PR 24-MAY-1999: 99US-0135629.
PR 25-MAY-1999: 99US-0136021.
PR 27-MAY-1999: 99US-0136392.
PR 28-MAY-1999: 99US-0136782.
PR 01-JUN-1999: 99US-0137222.
PR 03-JUN-1999: 99US-0137528.
PR 04-JUN-1999: 99US-0137502.
PR 07-JUN-1999: 99US-0137724.
PR 08-JUN-1999: 99US-0138094.
PR 10-JUN-1999: 99US-0138540.
PR 10-JUN-1999: 99US-0138847.
PR 14-JUN-1999: 99US-0139119.
PR 16-JUN-1999: 99US-0139432.
PR 16-JUN-1999: 99US-0139453.
PR 17-JUN-1999: 99US-0139453.
PR 18-JUN-1999: 99US-0139454.
PR 18-JUN-1999: 99US-0139455.
PR 18-JUN-1999: 99US-0139456.
PR 18-JUN-1999: 99US-0139457.
PR 18-JUN-1999: 99US-0139458.
PR 18-JUN-1999: 99US-0139459.
PR 18-JUN-1999: 99US-0139460.
PR 18-JUN-1999: 99US-0139461.
PR 18-JUN-1999: 99US-0139462.
PR 18-JUN-1999: 99US-0139463.
PR 18-JUN-1999: 99US-0139750.
PR 18-JUN-1999: 99US-0139750.
PR 18-JUN-1999: 99US-0139750.
PR 21-JUN-1999: 99US-0139750.
PR 22-JUN-1999: 99US-0139750.
PR 23-JUN-1999: 99US-0139750.
PR 23-JUN-1999: 99US-0139750.
PR 24-JUN-1999: 99US-0140082.
PR 28-JUN-1999: 99US-0140082.
PR 29-JUN-1999: 99US-0140091.
PR 30-JUN-1999: 99US-0141247.
PR 01-JUL-1999: 99US-0141842.
PR 02-JUL-1999: 99US-0142154.
PR 06-JUL-1999: 99US-0142055.
PR 08-JUL-1999: 99US-0142055.
PR 09-JUL-1999: 99US-0142803.
PR 09-JUL-1999: 99US-0142920.
PR 12-JUL-1999: 99US-0142977.
PR 13-JUL-1999: 99US-0143542.
PR 14-JUL-1999: 99US-0143624.
PR 15-JUL-1999: 99US-0144005.
PR 16-JUL-1999: 99US-0144085.
PR 16-JUL-1999: 99US-0144086.
PR 16-JUL-1999: 99US-0144325.
PR 19-JUL-1999: 99US-0144331.
PR 19-JUL-1999: 99US-0144332.
PR 19-JUL-1999: 99US-0144333.
PR 19-JUL-1999: 99US-0144334.
PR 19-JUL-1999: 99US-0144335.
PR 20-JUL-1999: 99US-0144352.
PR 20-JUL-1999: 99US-0144632.
PR 20-JUL-1999: 99US-0144884.
PR 21-JUL-1999: 99US-0144814.
PR 21-JUL-1999: 99US-0145086.
PR 21-JUL-1999: 99US-0145088.
PR 22-JUL-1999: 99US-0145085.
PR 22-JUL-1999: 99US-0145087.
PR 22-JUL-1999: 99US-0145089.
```

PR	22-JUL-1999	9905-0145192
PR	22-JUL-1999	9905-0145145
PR	22-JUL-1999	9905-0145218
PR	22-JUL-1999	9905-0145224
PR	26-JUL-1999	9905-0145276
PR	27-JUL-1999	9905-0145918
PR	27-JUL-1999	9905-0145918
PR	27-JUL-1999	9905-0145919
PR	28-JUL-1999	9905-0145951
PR	02-AUG-1999	9905-0146386
PR	02-AUG-1999	9905-0146388
PR	02-AUG-1999	9905-0146389
PR	03-AUG-1999	9905-0147038
PR	03-AUG-1999	9905-0147204
PR	04-AUG-1999	9905-0147302
PR	05-AUG-1999	9905-0147192
PR	05-AUG-1999	9905-0147260
PR	05-AUG-1999	9905-0147303
PR	06-AUG-1999	9905-0147416
PR	09-AUG-1999	9905-0147493
PR	09-AUG-1999	9905-0147935
PR	10-AUG-1999	9905-0148171
PR	11-AUG-1999	9905-0148319
PR	12-AUG-1999	9905-0148341
PR	13-AUG-1999	9905-0148365
PR	13-AUG-1999	9905-0148684
PR	15-AUG-1999	9905-0149368
PR	17-AUG-1999	9905-0149175
PR	18-AUG-1999	9905-0149426
PR	20-AUG-1999	9905-0149722
PR	20-AUG-1999	9905-0149723
PR	20-AUG-1999	9905-0149929
PR	23-AUG-1999	9905-0149930
PR	25-AUG-1999	9905-0150065
PR	25-AUG-1999	9905-0150084
PR	26-AUG-1999	9905-0151065
PR	27-AUG-1999	9905-0151066
PR	27-AUG-1999	9905-0151080
PR	30-AUG-1999	9905-0151303
PR	31-AUG-1999	9905-0151438
PR	01-SEP-1999	9905-0151930
PR	01-SEP-1999	9905-0152363
PR	07-SEP-1999	9905-0153070
PR	10-SEP-1999	9905-0153758
PR	13-SEP-1999	9905-0154018
PR	15-SEP-1999	9905-0154039
PR	16-SEP-1999	9905-0154779
PR	22-SEP-1999	9905-0155131
PR	22-SEP-1999	9905-0155186
PR	24-SEP-1999	9905-0155659
PR	28-SEP-1999	9905-0156458
PR	20-SEP-1999	9905-0156596
PR	04-OCT-1999	9905-0157117
PR	05-OCT-1999	9905-0157865
PR	06-OCT-1999	9905-0157853
PR	07-OCT-1999	9905-0158029
PR	08-OCT-1999	9905-0158369
PR	12-OCT-1999	9905-0159293
PR	13-OCT-1999	9905-0159294
PR	13-OCT-1999	9905-0159325
PR	13-OCT-1999	9905-0159329
PR	14-OCT-1999	9905-0159330
PR	14-OCT-1999	9905-0159331
PR	14-OCT-1999	9905-0159637
PR	14-OCT-1999	9905-0159638
PR	18-OCT-1999	9905-0160741
PR	21-OCT-1999	9905-0160747
PR	21-OCT-1999	9905-0160768
PR	21-OCT-1999	9905-0160770
PR	21-OCT-1999	9905-0160814

PR 21-OCT-1999: 9905-0160815.  
PR 22-OCT-1999: 9905-0160730.  
PR 22-OCT-1999: 9905-0160981.  
PR 22-OCT-1999: 9905-0160989.  
PR 25-OCT-1999: 9905-0161404.  
PR 25-OCT-1999: 9905-0161405.  
PR 25-OCT-1999: 9905-0161406.  
PR 25-OCT-1999: 9905-0161359.  
PR 26-OCT-1999: 9905-0161360.  
PR 26-OCT-1999: 9905-0161361.  
PR 26-OCT-1999: 9905-0161920.  
PR 28-OCT-1999: 9905-0161920.  
PR 28-OCT-1999: 9905-0161992.  
PR 28-OCT-1999: 9905-0161993.  
PR 29-OCT-1999: 9905-0162142.

[illegible]

## RESULT 11

AAV19497 standard; DNA; 585 BP.

AC AAV19497;

DT 18-SEP-1998 (first entry)

Cytochrome P4501Id6 encoding DNA 847-1431.

Human; cytochrome P4501d6; liver/kidney microsome; LKM-1; antibody;

KW Human; cytolome F400100; HIV/AIDS; autoimmunohepatitis type II; AIH type II; hepatitis C virus; ss.

OS Hepatitis C virus.

aa	Key	Level 100/Qualifiers
HH		

FT 1.585

FT	product	"cytochrome P4501d6 fragment"
FT	/transl_except-	(pos: 316..318,aa:Xaa)
FT		

14  
FT /note- "Xad is unspecified"

PN JP10101698-A.

PD 21-APR-1998.

24-SEP-1996; 96JP-0273055-PF

AA 24-SEP-1996; 96JP-0273055.  
PR

PA (HOKE-) HOKEN KAGAKU KENKYUSHO KK.

DR WPI: 1998-292131/26.

DR P-PSDB; AAW44870.

PT Amino acid sequence for liver/kidney microsome (LKM) 1 antibody - to

human auto-immuno-hepatitis (AIH) type II and hepatitis C

PS Claim 3; Fig 3; 9pp; Japanese.

<sup>aa</sup> CC The present sequence encodes a protein fragment of cytochrome P450IIdb







PA (HAYA/) HAYASHI K.  
PA (SUMO) SUMITOMO CHEM CO LTD.

XX  
PI Hayashi K, Kaneko H, Komai K, Nakatsuka I, Sakaki T,

PI Yabusaki Y;

WPI; 1995-116991/16.

DR P-PSDB; AAR72378.

PT Evaluation of safety of a chemical cpd. - using recombinant yeast  
PT expressing human cytochrome p450 and a yeast NADPH-P450 reductase

PS Examples; page 95-97; 124pp; English.

The nucleotide sequence of the cDNA coding region for the human auxiliary cytochrome P450 species 206 variant 3. The gene contains a change at base 1457 from C to G as compared to the wild type sequence (AA087723). This changes the amino acid residue 296 from Thr to Ser. The cDNA was amplified by PCR using the primers AA087733-6. The product was cloned into the yeast expression vectors pANSh or pANR+ to produce the vectors p206 v. variant 3 for the expression of the cytochrome P450 or p206R variant 3 for co-expression with the yeast NADPH-P450 reductase. The vectors are used in a method for evaluating the safety of a chemical compound by reacting the chemical compound with recombinantly produced human cytochrome P450 molecular species 1A2 (AA087716), 2C9 (AA087715), 2E1 (AA087717), or 3A4 (AA087717) or their auxiliary species and variants (AA087718-32), and yeast NADPH-P450 reductase, either as a fused protein or in cell extracts, and analysing the resulting metabolite to assess the safety of the chemical compound. The method is useful for determining whether the chemical compound, or its metabolite, will be converted into a carcinogenic or mutagenic form through metabolism in the liver.

Sequence 1494 BP; 248 A; 507 C; 447 G; 292 T; 0 other

Query Match	Score	DB ID	Length
6.58	32.8	16	1494
61.98	pred	NO. 3	

Best local Similarity	61.98;	Pred. No.	37
-----------------------	--------	-----------	----

Matches	52;	Conservative	0;	Mismatches	32;	Indels	0;	Gaps	0
---------	-----	--------------	----	------------	-----	--------	----	------	---

**OY**

345 tcdgaccatcctccgcgatcgaggactacagtccttctgtctctggcatq 404  
||||| || | |||| | |||| | || | ||  
620 ttattaaccaacctgaacaaacucctugucccttgagaactctctcacatgccatacccagaty 979

Db 920 tggltgaccaccttcgaaccatcgtgcctcggtgccctccctcccaagacccccttctttct

QY 405 tgcgaggaqacgcgaggaqcaqa 428

Db 980 tgcagcgcctgllccaacagga 1003

Search completed: March 4, 2002, 05:24:09  
Job time: 17678 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OR nucleic - nucleic search, using sw model

Run on: March 4, 2002, 05:12:49 ; Search time 226.99 seconds  
(without alignments)  
499,870 Million cell updates/sec

Title: US-08-956-991-9\_Copy\_500\_1000

Perfect score: 1 accttcgcgcagcagcagata.....aacacacagcagcagatac 501

Sequence: 1 accttcgcgcagcagcagata.....aacacacagcagcagatac 501

Scoring table: IDENTITY\_MMC  
Gapop 10.0, Gapext 1.0

Searched: 351203 seqs, 113218999 residues

1 number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%  
Listing first 45 summaries

Database :  
1: /cgn2\_6/ptodata/2/1na/5A.COMB.seq.\*  
2: /cgn2\_6/ptodata/2/1na/5B.COMB.seq.\*  
3: /cgn2\_6/ptodata/2/1na/6A.COMB.seq.\*  
4: /cgn2\_6/ptodata/2/1na/6B.COMB.seq.\*  
5: /cgn2\_6/ptodata/2/1na/PCPUS.COMB.seq.\*  
6: /cgn2\_6/ptodata/2/1na/back11test.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	length	DB ID	Description
1	33.8	6.7	1171	1	US-08-336-257A-1
2	33.8	6.7	1171	6	Patent No. 5386025-1
3	32.6	6.5	1396	1	US-08-123-161A-11
4	32.6	6.5	1396	1	US-08-484-278-11
5	32.6	6.5	1436	4	US-08-910-864-11
6	31.8	6.3	533	6	5482709-5
7	31.8	6.3	543	6	5273901-6
8	31.2	6.2	1569	2	US-08-145-658D-23
9	30.8	6.1	1050	1	US-08-599-252-81
10	30.8	6.1	1050	1	US-08-436-074-54
11	30.8	6.1	1050	5	PCT-US96-06352-81
12	30.8	6.1	1050	5	PCT-US96-06583-81
13	30.6	6.1	1501	2	US-08-784-651-3
14	30.4	6.1	1568	2	US-08-145-658D-24
15	30.4	6.1	1701	2	US-07-863-169A-13
16	30.4	6.1	1701	2	US-08-429-964-2
17	30.4	6.1	1701	2	US-07-935-087-2
18	30.4	6.1	1701	5	PCT-US93-08062-20
19	30.4	6.1	1701	5	PCT-US93-08062-20
20	30.2	6.0	1568	2	US-08-469-802B-1
21	30	6.0	3366	2	US-08-267-803A-1
22	30	6.0	10660	2	US-08-267-803B-8
23	30	6.0	10660	4	US-09-041-886-16
24	29.8	5.9	1483	6	5508199-7
25	29.8	5.9	3489	2	US-08-728-123A-1
26	29.8	5.9	3489	2	US-08-770-379-20
27	29.8	5.9	32207	2	US-08-770-379-20

28	29.8	5.9	32207	4	US-08-757-669A-20	Sequence 20, Appl
29	29.4	5.9	1566	2	US-08-145-658D-22	Sequence 22, Appl
30	29.4	5.9	1571	2	US-08-145-658D-21	Sequence 21, Appl
31	29.4	5.9	252	4	US-09-115-446-1	Sequence 1, Appl
32	29.4	5.9	2526	4	US-09-115-446-5	Sequence 5, Appl
33	29.4	5.9	5822	3	US-08-899-595-4	Sequence 4, Appl
34	29.4	5.9	5822	3	US-08-899-595-5	Sequence 5, Appl
35	29.2	5.8	435	3	US-08-301-162-9	Sequence 9, Appl
36	29.2	5.8	623	4	US-09-043-303-5	Sequence 7, Appl
37	29.2	5.8	1140	1	US-08-424-268-7	Sequence 7, Appl
38	29.2	5.8	1140	5	PCT-US93-10442-7	Sequence 6, Appl
39	29.2	5.8	1664	1	US-07-863-169A-6	Sequence 6, Appl
40	29.2	5.8	1664	2	US-08-429-964-6	Sequence 6, Appl
41	29.2	5.8	1664	4	US-07-935-087-6	Sequence 6, Appl
42	29.2	5.8	1664	5	PCT-US93-08062-6	Sequence 6, Appl
43	29.2	5.8	2546	1	US-08-424-268-19	Sequence 19, Appl
44	29.2	5.8	2546	5	PCT-US93-10442-19	Sequence 19, Appl
45	29.2	5.8	291	1	US-08-188-277B-6	Sequence 6, Appl

## ALIGNMENTS

RESULT 1  
US-08-336-257A-1  
Sequence 1, Application US/08336257A  
Patent No. 5726045  
GENERAL INFORMATION:  
APPLICANT: Jay, Scott D.  
APPLICANT: Ellis, Steven H.  
APPLICANT: Harpold, Michael M.  
APPLICANT: Campbell, Kevin P.  
TITLE OF INVENTION: CALCIUM CHANNEL COMPOSITIONS AND METHODS  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Brown, Martin, Haller & McClain  
STREET: 1660 Union Street  
CITY: San Diego  
STATE: CA  
COUNTRY: USA  
ZIP: 92101-2926  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/336,257A  
FILING DATE: 07-Nov-1994  
CLASSIFICATION: 433  
ATTORNEY/AGENT INFORMATION:  
NAME: Seddiqui, Syedul Karim L.  
REGISTRATION NUMBER: 35,779  
REFERENCE/BOOK NUMBER: 54898  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 238-0062  
TELEFAX: (619) 238-0062  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1171 base pairs  
TYPE: nucleic acid  
STRANDINESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 49..717  
OTHER INFORMATION: /product- "Gamma subunit of animal  
OTHER INFORMATION: calcium channel"  
US-08-336-257A-1  
Query Match 6.7%: Score 33.8; DB 1: length 1171;



TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 4..1164  
US-08-483-278-11

Query Match 6.5%; Score 32.6; DB 1; Length 1396;  
Best Local Similarity 58.9%; Pred. No. 0.61;  
Matches 56; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 444 ctcgtaaccatctcctgcaatctcttcaatctcttctggtctctgctggtt 403  
DB 874 ctcgtaaccatctcctgcaatctcttcaatctcttctggtctctggtt 933  
QY 404 gtcgaag 438  
DB 934 gtcgaag 968

RESULT 5  
US-08-910-864-11  
Sequence 11, Application US/08910864  
Patent No. 6280531  
GENERAL INFORMATION:  
APPLICANT: SAKAMOTO, AII  
APPLICANT: HANAKO, FUMIO  
TITLE OF INVENTION: METHOD FOR SPECIFICALLY AMPLIFYING A CDNA OF AN EXTREMELY  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FISH & RICHARDSON P.C.  
STREET: 4225 EXECUTIVE SQUARE, SUITE 1400  
CITY: LA JOLLA  
STATE: CA  
COUNTRY: USA  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatcomLin Release #1.0, Version #1.40 (RPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/910,864  
FILING DATE: 13-AUG-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 216506/1996  
FILING DATE: 16-AUG-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: BAILE, LISA A.  
REGISTRATION NUMBER: 48,347  
REFERENCE/DOCKET NUMBER: 07898/017001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-678-5070  
TELEFAX: 619-678-5099  
INFORMATION FOR SEQ. ID NO.: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 143; base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA to mRNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 44..1204  
US-08-910-864-11

Query Match 6.5%; Score 32.6; DB 1; Length 1396;

Best Local Similarity 58.9%; Pred. No. 0.62;  
Matches 56; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 344 ctcgtaaccatctcctgcaatctcttcaatctcttctggtctctggtt 403  
DB 914 ctcgtaaccatctcctgcaatctcttcaatctcttctggtctctggtt 973  
QY 404 gtcgaag 438  
DB 974 gtcgaag 1008

RESULT 6  
5482709-5/c  
Patent No. 5482709  
APPLICANT: JACKSON, JAMES W.; STRAUSBERG, ROBERT L.; WILSON,  
SUSAN D.; POPE, SHARON H.; STRAUSBERG, SUSAN L.; ROPE, MICHAEL D.;  
AUGUSTINE, PATRICIA C.; DANFORTH, HARRY D.  
TITLE OF INVENTION: ERMERIA ANTIGENIC COMPOSITION WHICH  
ELICITS ANTIBODIES AGAINST AVIAN COCCIDIOSIS  
NUMBER OF SEQUENCES: 10  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/148,432  
FILING DATE: 08-NOV-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 581,693  
FILING DATE: 12-SEP-1990  
APPLICATION NUMBER: 215,162  
FILING DATE: 05-JUL-1989  
APPLICATION NUMBER: 746,420  
FILING DATE: 19-JUN-1985  
APPLICATION NUMBER: 627,811  
FILING DATE: 05-JUL-1984  
SEQ ID NO: 5  
LENGTH: 544  
5482709-5

Query Match 6.4%; Score 31.8; DB 6; Length 533;  
Best Local Similarity 56.1%; Pred. No. 0.65;  
Matches 60; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 345 ctcgaagatctctgcaatctcttcaatctcttctggtctctggtt 394  
DB 345 ctcgaagatctctgcaatctcttcaatctcttctggtctctggtt 286  
QY 435 ctcgtaaccatctcctgcaatctcttcaatctcttctggtctctggtt 441  
DB 285 ctcgtaaccatctcctgcaatctcttcaatctcttctggtctctggtt 239

RESULT 7  
5273901-6/c  
Patent No. 5273901  
APPLICANT: JACKSON, JAMES W.; STRAUSBERG, ROBERT L.; WILSON,  
SUSAN D.; POPE, SHARON H.; STRAUSBERG, SUSAN L.; ROPE, MICHAEL D.;  
AUGUSTINE, PATRICIA C.; DANFORTH, HARRY D.  
TITLE OF INVENTION: GENETICALLY ENGINEERED COCCIDIOSIS  
SPOROZONTE 21.5 K8 ANTIGEN, AC-68  
NUMBER OF SEQUENCES: 11  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/01/581,693  
FILING DATE: 12-SEP-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 215,162  
FILING DATE: 05-JUL-1988  
APPLICATION NUMBER: 746,520  
FILING DATE: 19-JUN-1985  
APPLICATION NUMBER: 627,811  
FILING DATE: 05-JUL-1984  
SEQ ID NO: 6  
LENGTH: 543  
5273901-6



```

1  APPLICANT: WOLFE, ROGER K.
2  TITLE OF INVENTION: METHOD TO DIAGNOSE HEREDITARY
3  TITLE OF INVENTION: HEMOCHROMATOSIS
4  NUMBER OF SEQUENCES: 124
5  CORRESPONDENT'S ADDRESS:
6  ADDRESSEE: MORRISON & FOEKSTER
7  STREET: 2000 Pennsylvania Ave., N.W., Suite 5500
8  CITY: Washington
9  STATE: DC
10 COUNTRY: USA
11 ZIP: 20006-1888
12
13 COMPUTER READABLE FORM:
14 MEDIUM TYPE: floppy disk
15 COMPUTER: IBM PC compatible
16 OPERATING SYSTEM: PC-DOS/MS-DOS
17 SOFTWARE: Patent In Release #1.0, Version #1.30
18
19 CURRENT APPLICATION DATA:
20 APPLICATION NUMBER: PCT/US96/06352
21 FILING DATE:
22 CLASSIFICATION:
23 PRIOR APPLICATION DATA:
24 APPLICATION NUMBER: US 08/599,252
25 FILING DATE: 09-FEB-1996
26 ATTORNEY/AGENT INFORMATION:
27 NAME: MORASITICH, KATE H.
28 REGISTRATION NUMBER: 29,959
29 REFERENCE/BOOKET NUMBER: 9054-0001.21
30 TELECOMMUNICATION INFORMATION:
31 TELEPHONE: (202) 887-1500
32 TELEFAX: (202) 887-9763
33 TELE: 90-4030
34
35 INFORMATION FOR SEQ ID NO: 81:
36 SEQUENCE CHARACTERISTICS:
37 LENGTH: 1050 base pairs
38 TYPE: nucleic acid
39 STRANDEDNESS: single
40 TOPOLOGY: linear
41
42 PCT-US96-06352-81
43
44 Query Match 0.16; Score 30.8; DB 5; Length 1050;
45 Best Local Similarity 54.88; Freq. No. 2;
46 Matches 65; Conserved 0; Mismatches 57; Indels 0; Gaps 0.
47
48 QY 334 gctcaagatcctgagacatctctgcagctcctgcagtgaggtcctcactctcttgct 393
49 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
50 DB 301 gcccaagctatttattatagtttcaacatctgctctgcttgcgtgggagcttggccctctgct 360
51 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
52 QY 394 tctctgattctctgagagagagagagagagagagagagagagagagagagagagagag 453
53 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
54 DB 361 tctctgattctctgagagagagagagagagagagagagagagagagagagagagagag 420
55 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
56 QY 454 gaa 455
57 DB 421 GA 422
58
59 RESULT 12
60 PCT-US96-06584-81
61 Sequence 81, Application PCT/US9606584
62 GENERAL INFORMATION:
63 APPLICANT: DRAVINA, DENNIS F.
64 APPLICANT: FEDER, JOHN N.
65 APPLICANT: KIMBLE, ANDREAS
66 APPLICANT: KIMBLE, BRUCE E.
67 APPLICANT: THOMAS, WINSTON J.
68 APPLICANT: WOLFE, ROGER K.
69 TITLE OF INVENTION: METHOD TO DIAGNOSE HEREDITARY
70 TITLE OF INVENTION: HEMOCHROMATOSIS
71 NUMBER OF SEQUENCES: 124
72 CORRESPONDENCE ADDRESS:
73 ADDRESSEE: MORRISON & FOEKSTER
74 STREET: 2000 Pennsylvania Ave., Suite 5500
75

```



CITY: Washington  
 STATE: DC  
 COUNTRY: USA  
 ZIP: 20006-1888  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: PCT/US96/06583  
 FILING DATE:  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/594,252  
 FILING DATE: 09-FEB-1996  
 ATTORNEY/AGENT INFORMATION:  
 NAME: MORASHIGE, KATE H.  
 REGISTRATION NUMBER: 29,959  
 REFERENCE/DOCKET NUMBER: 9053-0001.21  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (202) 887-1500  
 TELEFAX: (202) 887-0763  
 TELFAX: 90-4010  
 INFORMATION FOR SEQ ID NO: 81:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1050 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: Single  
 TOPOLOGY: Linear

101-0596-06583-81

```

Query Match: 6-18; Score 30.8; Dh 5; Length 1050;
Best Local Similarity 53.3%; Prot. No. 2;
Matches 65; Conservative 0; Mismatches 57; Indels 0; Gaps 0

Dh 334 gctcaaatccctcgagacatctctctgacccctgacgcaatcttaactgctcttgact 393
      11 1111 1 1111 1 1111 1 1111 1 11 11 11 1111
Dh 301 GCCCAAGCTATTGCTGAGCTTGGCAAGTGTGGCTGTGGGTGGAGTGTGGT 360
      11 1111 1 1111 1 1111 1 1111 1 11 11 1111 11
Dh 361 TCTCTGGAGCTCTGACAGCTCTGCTCTGCTGGCTTACCTCTGCTGACAGAGAAAGATGAA 420
      11 11 11 11 11 11 11 11 11 11 11 11 11
Dh 454 ga 455
      11
Dh 421 GA 422

RESULT 13
US-08-784-651-4/c
Sequence 3, Application US/08784651
Patent No. 5821102
GENERAL INFORMATION:
APPLICANT: Betka, Randy
APPLICANT: Boomlathan, Karuppan
APPLICANT: Sandali, Thomas
TITLE OF INVENTION: Nucleic Acids Encoding Polypeptides
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESS: No. 58211020 No. 5821102disk of No. 5821102th America, Inc.
STREET: 405 Lexington Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10174
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSB0 for Windows Version 2.0
CURRENT APPLICATION DATA:

```

1 APPLICATION NUMBER: US/08/784,651  
 2 FILING DATE: 21-JAN-1997  
 3 CLASSIFICATION: A35  
 4 ATTORNEY/AGENT INFORMATION:  
 5 NAME: Agtis, Cheryl H.  
 6 REGISTRATION NUMBER: 34,086  
 7 RESEARCH/DOCKET NUMBER: 4608,200  
 8 TELECOMMUNICATION INFORMATION:  
 9 TELEPHONE: 212-867-0123  
 10 TELEFAX: 212-878-9655  
 11 INFORMATION FOR SEQ ID NO: 1:  
 12 SEQUENCE CHARACTERISTICS:  
 13 LENGTH: 2149 base pairs  
 14 TYPE: nucleic acid  
 15 STRANDEDNESS: single  
 16 TOPOLOGY: linear  
 17 MOLECULE TYPE: Genomic DNA  
 18 US-08-784-651-1

[illegible]

1 RESULT 1  
2 US-08-145-6580-24  
3 Sequence 24, Application US/081456580  
4 Patent No. 5981174  
5 GENERAL INFORMATION:  
6 APPLICANT: Wolff, Charles R.  
7 APPLICANT: Miles, John S.  
8 APPLICANT: Spier, Nicole K.  
9 APPLICANT: Goodall, Alan  
10 TITLE OF INVENTION: GEN. ITC. ASSAY  
11 NUMBER OF SEQUENCES: 2  
12 CORRESPONDENCE ADDRESS:  
13 ADDRESSEE: DICKSTEIN, SHAPIRO, MORIN & OSHINSKY LLP  
14 STREET: 2101 E. Street N.W.  
15 CITY: Washington  
16 STATE: DC  
17 COUNTRY: USA  
18 ZIP: 20037  
19 COMPUTER READABLE FORM:  
20 MEDIUM TYPE: Floppy disk  
21 COMPUTER: IBM PC compatible  
22 OPERATING SYSTEM: PC-DOS/MS-DOS  
23 SOFTWARE: Patent In Release #1.0, Version #1.30  
24 CURRENT APPLICATION DATA:  
25 APPLICATION NUMBER: US/08/145,6580  
26 FILING DATE: 04-MAY-1994  
27 CLASSIFICATION: A35  
28 PRIOR APPLICATION DATA:  
29 APPLICATION NUMBER: GB 9001181.8  
30 FILING DATE: 18-JAN-1990  
31 PRIOR APPLICATION DATA:  
32 APPLICATION NUMBER: US 07/742,223  
33 FILING DATE: 18-JUL-1991  
34 ATTORNEY/AGENT INFORMATION:  
35 NAME: Brady, Jr., James W.  
36 REGISTRATION NUMBER: 42,115  
37 REFERENCE/DOCKET NUMBER: EB280.017  
38 TELECOMMUNICATION INFORMATION:  
39 TELEPHONE: 202-785-9700  
40 TELEFAX: 202-887-0689

```

1 INFORMATION FOR SEQ ID NO: 24
2
3 SEQUENCE CHARACTERISTICS:
4
5 LENGTH: 1501 base pairs
6
7 type: nucleic acid
8
9 STRANDEDNESS: single
10
11 TOPOLOGY: linear
12
13 MOLECULE TYPE: cDNA to mRNA
14
15 HYPOTHETICAL: NO
16
17 ANTI-SENSE: NO
18
19 US-08-145-658D-24

```

Query Match	6.18;	Score 30.4;	DB 2;	Length 1501;
Best local Similarity	61.28;	Pred. No. 3.4;		
Matches 49;	Conservative 0;	Mismatches 31;	Indels 0;	Gaps 0;

[illegible]

RESULT 15-6580-13  
 US-08-145-6580-13  
 Sequence 13, Application US/08145658D  
 Patent No. 5981174  
 GENERAL INFORMATION:  
 APPLICANT: Wolf, Charles R.  
 APPLICANT: Miles, John S.  
 APPLICANT: Spurr, Nigel K.  
 APPLICANT: Coughlin, Alan C.  
 TITLE OF INVENTION: GENETIC ASSAY  
 NUMBER OF SEQUENCES: 25  
 CORRESPONDENCE ADDRESS:  
 ADDRESSER: DICKSTEIN, SHAPIRO, MORIN & OSHINSKY LLP  
 STREET: 2101 L Street N.W.  
 CITY: Washington  
 STATE: DC  
 COUNTRY: USA  
 ZIP: 20037  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent in Release #1.0, Version #1.10  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/145,658D  
 FILING DATE: 04-NOV-1993  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: GR 9001181.8  
 FILING DATE: 18-JAN-1990  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/732,224  
 FILING DATE: 18-JUL-1991  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Brady, Jr., James W.  
 REGISTRATION NUMBER: 32,115  
 REFERENCE/DOCKET NUMBER: E8280.017  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 202-765-9700  
 TELEFAX: 202-867-0689  
 INFORMATION FOR SEQ ID NO: 13:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1566 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA to mRNA  
 HYPOTHETICAL: NO

; ANTI-SENSE: NO  
 ; ORIGINAL SOURCE:  
 ; ORGANISM: Homo sapiens  
 US-08-145-658D-13

Query Match	6.18	Score 30.4	DB 2	Length 1566
Best Local Similarity	63.99	Pred. No. 3.4		
Matches	46	Conservative	0	Mismatches 26
				Indels 0
				Gaps 0

OY		345	lclgagaaacatctccgggaactcgcgtttagtcctcgcaagttag	404
Dδ		920	tggcacattacctaccataaaccttgactttggaggcttcccc	
OY		405	lgtcgagaqanaac	416
Dβ		980	tgcacattctctctc	991

Search completed: March 4, 2002, 05:13:06  
Job time: 17034 sec